THE EFFECTS OF ACCOUNTABILITY ON INFORMATION SEARCH, EVALUATION AND INTEGRATION IN MULTIATTRIBUTE DECISION MAKING

genehmigte Dissertation

zur Erlangung des akademischen Grades eines Doktors der Philosophie (Dr. phil.)

der Technischen Universität Darmstadt

Fachbereich 03: Humanwissenschaften

vorgelegt von Dipl.-Psych. Stefanie Schmeer, geboren in Darmstadt

eingereicht am 10.09.01

mündliche Prüfung am 11.04.02

Referenten: Prof. Dr. Katrin Borcherding

Prof. Dr. Bruno Rüttinger

D17

Darmstadt, im Dezember 2002

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor, Prof. Dr. Katrin Borcherding, for her absolutely invaluable practical and moral support, which she has continued to provide even after I left the Technical University of Darmstadt, and which has gone far beyond any call of duty.

I am also very grateful to all student research assistants who helped me with the onerous tasks of data collection and entry; for their help with a pilot experiment, Petra Schreiber, Andreas Schubert and Jürgen Junker, for their help with the first experiment, Heike Kollross and Claudia Toll, and for their help with the second experiment, Meike Pflanz, Tobias Ley and Tanja Merle. I am particularly indebted to Claudia Toll, who created the Word macros that made the nightmare of creating individually randomised questionnaires bearable.

Many thanks also to Petra Knötzinger, who introduced me to the world of CASE tools and whose expertise as an IT consultant was vital for creating a realistic decision scenario for the second study. I was also supported in this task by Bettina Kühnapfel from the Computer Science department at the Technical University of Darmstadt, who very kindly took the time to explain principles of software engineering to a lay-person and encouraged her students to participate in my experiment.

I would furthermore like to thank the Computer Science department at the Technical University of Darmstadt for allowing me to use one of their rooms for data collection and Kai Flade, who helped to recruit participants for this study. Many thanks also to all people who agreed to participate in my experiments, sometimes even without accepting any reward.

I am particularly grateful to all my friends and family, whom I have not been able to spend as much time with as I would have liked, for not giving up on me, despite my frequent and long absences. Many thanks to Emma Palmer, Jackie Andrade, Chris MacDonald and Norma Kelly, Jürgen Grafe, Berthold Kuttich, Dorrit and Peter Linnell, and Annelise and Karl Hartig. Thank you also to my friends from the University of Hertfordshire and to Lis Meyland-Smith for their moral support. Although she cannot see it anymore, I would also like to mention my grand aunt, Margarete Völker, at this point, who was always very supportive of me.

A special, big thankyou goes to my mother, Rosemarie Schmeer, who has supported me throughout this time in the way she knows best, by looking after me and spoiling me when I have come to visit.

Finally, to K.J.L. This has been a very long, difficult and costly time, also for you. Thank you.

ZUSAMMENFASSUNG

Die vorliegende Arbeit befasst sich mit dem Einfluss von Rechtfertigungsdruck auf die Suche, Bewertung und Integration von Informationen bei der multiattributiven Entscheidungsfindung. Dabei wurden klassische sozialpsychologische und entscheidungsanalytische Ansätze miteinander kombiniert. Rechtfertigungsdruck wurde als eine motivierende Variable angesehen, die bestimmte Ziele in die Informationsverarbeitung des Entscheiders einführt. Welche Ziele das im einzelnen sind, hängt von situativen Merkmalen des Rechtfertigungsdrucks ab. Es wurde gezeigt, dass Rechtfertigungsdruck im wesentlichen aus zwei Komponenten besteht, (1) dem Gefühl persönlicher Verantwortlichkeit für eventuelle negative Konsequenzen des eigenen Verhaltens und (2) der Bewertungserwartung des eigenen Verhaltens durch eine „Öffentlichkeit“ (externes Publikum). Die zweite Komponente ist jedoch nur dann vorhanden, wenn erwartet wird, dass der Prozess der Entscheidungsfindung von einer außenstehenden Person beobachtet wird bzw. dieser gegenüber gerechtfertigt werden muss (externer Rechtfertigungsdruck). Bei internem Rechtfertigungsdruck, d.h. einem Verlangen, das eigene Verhalten sich selbst gegenüber rechtfertigen zu können, ist Bewertungserwartung minimal, da es kein externes Publikum gibt.

Es wird ein Prozessmodel der Rechtfertigung (process model of accountability, PMA) entwickelt. Dessen Hauptannahme ist es, dass unter Rechtfertigungsdruck der Entscheidungsprozess sorgfältiger und komplexer, aber gleichzeitig stärker verfälscht abläuft. Rechtfertigungsdruck hat deshalb nicht die ausschließlich positive Wirkung, die ihm gemeinhin zugesprochen wird. Das Prozessmodel des Rechtfertigung nimmt weiterhin an, dass Hinweise auf persönliche Verantwortlichkeit den Wunsch nach Genauigkeit erhöhen. Letztere versuchen Entscheider dadurch zu erreichen, dass sie die Tiefe ihrer Informationssuche und Komplexität ihrer Informationsverarbeitung erhöhen. Dahingegen erzeugt Bewertungserwartung sogenannte Richtungsziele, d.h. Ziele, die den Entscheidungsprozess in eine bestimmte Richtung verfälschen, z.B. in Richtung einer Anpassung an die Präferenzen eines externen Publikums bzw., im Falle der erneuten Entscheidung, in Richtung auf eine Bestätigung der vorherigen Entscheidung.

Zur Überprüfung einiger zentraler, aus dem Prozessmodell abgeleiteter Vorhersagen wurden zwei Experimente mit folgender Versuchsanordnung durchgeführt. Die Versuchsteilnehmer sollten nach einer ersten Suche von Informationen eine vorläufige Entscheidung fällen. Bevor eine endgültige Entscheidung getroffen werden musste, konnten sie zusätzliche Informationen aus einer zweiten Informationsmenge einholen. Die für die zweite Informationssuche bereitgestellten Informationen waren hochgradig redundant und erlaubten es somit, zu überprüfen, ob die Informationssuche nach einer vorläufigen Entscheidung zugunsten dieser verfälscht wird. Weiterhin wurden mehrmals während des Entscheidungsprozesses Bewertungen der Attraktivität einzelner Aspekte von Alternativen sowie die Wichtigkeit dieser Aspekte für die Entscheidung abgefragt. Dies geschah, um zu überprüfen, ob sie vom Entscheider so verändert werden, dass sich der Attraktivitätsvorsprung der gewählten Alternative gegenüber den nicht gewählten Alternativen vergrößert. Letzteres wird von der Differenzierungs- und Konsolidierungstheorie (Svenson, 1992, 1996) vorhergesagt, welche Festingers (1957, 1964) Dissonanztheorie auf multiattributive Entscheidungssituationen übertragen und für solche spezifiziert hat.

Im ersten Experiment wurden die Effekte internen und externen Rechtfertigungsdrucks untersucht, der entweder vor oder nach einer ersten Entscheidung eingeführt wurde. Dabei wurden keine externen Präferenzen explizit gemacht. Es zeigte sich in Übereinstimmung zu den Vorhersagen des Prozessmodells, dass sowohl interner als auch externer Rechtfertigungsdruck in der Tendenz zu einer tieferen und komplexeren Informationssuche führte. Unter internem Rechtfertigungsdruck, welcher bereits vor einer ersten Entscheidung eingeführt wurde, blieb die Informationsverarbeitung auch nach einer vorläufigen Entscheidung relativ ausgewogen. Wurde interner Rechtfertigungsdruck jedoch erst nach einer vorläufigen Entscheidung eingeführt, verfälschten Entscheider ihre Informationssuche und Informationsbewertung stärker in Richtung der vorher gewählten Alternative als dies bei externem Rechtfertigungsdruck der Fall war. Verfälschungstendenzen bei der Informationsbewertung waren jedoch generell eher schwach und deuteten darauf hin, dass Personen unter externem Rechtfertigungsdruck die Attraktivitätsdifferenz zwischen ihrer gewählten Alternative und den nicht gewählten Alternativen reduzierten, wenn ihr Urteil offensichtlich war,

z.B. bei einer holistischen Bewertung der Güte von Alternativen, sie jedoch vergrößerten, wenn eine Verfälschungstendenz weniger deutlich war, z.B. wenn die Gesamtbewertung von Alternativen mit Hilfe der multiattributiven Nutzentheorie aus eindimensionalen Wert- und Gewichtsurteilen vorhergesagt wurde. Dies deutet darauf hin, dass Personen unter externem Rechtfertigungsdruck eine deutliche öffentliche Stellungnahme vermeiden wollen, wenn unklar ist, wie ein externes Publikum auf eine solche reagieren würde, und ihr Verhalten offensichtlich stark von Überlegungen geprägt ist, wie sie sich dem externen Publikum gegenüber positiv darstellen können.

Letztere Schlussfolgerung wurde auch von den Ergebnissen eines zweiten Experiments bestätigt, welches den Einfluss von Rechtfertigung gegenüber verschiedenen Interessengruppen auf Entscheidungsprozesse untersuchte. Die Versuchsteilnehmer wurden gebeten, für eine Firma eine Kaufentscheidung über ein Computerprogramm zu treffen und wurden entweder den späteren Benutzern dieses Programms oder dem Management der Firma gegenüber verantwortlich gemacht. Dies geschah entweder vor einer ersten Entscheidung oder erst, nachdem schon eine vorläufige Entscheidung getroffen worden war. Wie vom Prozessmodell der Rechtfertigung vorhergesagt, konnte eine Anpassung der Entscheidung an die Präferenzen der Interessengruppe, der gegenüber Versuchsteilnehmer verantwortlich gemacht worden waren, nur dann beobachtet werden, wenn Rechtfertigungsdruck schon vor einer ersten Entscheidung, nicht aber, wenn er erst nach einer ersten Entscheidung induziert worden war. Eine Anpassung an die Präferenzen des externen Publikums konnte insbesondere bei offensichtlichen Urteilen und weniger bei stärker versteckten Urteilen beobachtet werden. Bestand jedoch ein starker Konflikt zwischen den Präferenzen der Entscheider und den Präferenzen ihres externen Publikums, zeigten sowohl offene als auch versteckte Urteile Anzeichen einer Unterstützung der eigenen Präferenz. Dabei vergrößerten Entscheider die Attraktivitätsdifferenz zwischen ihrer gewählten Alternative und den nicht gewählten Alternativen und suchten Informationen, welche ihre gewählte Alternative im Vergleich zu den nicht gewählten Alternativen begünstigten. Gleichzeitig resultierte sowohl Rechtfertigungsdruck vor als auch nach einer ersten Entscheidung in einer tieferen und komplexeren Informationsverarbeitung im Vergleich zu einer Kontrollgruppe, bei der zu keinem Zeitpunkt Rechtfertigungsdruck erzeugt wurde. Dies

bestätigt, dass Rechtfertigungsdruck den Entscheidungsprozess sowohl sorgfältiger und komplexer macht als auch stärker verfälscht. Dabei scheint die erhöhte Tiefe der Verarbeitung Personen darin zu unterstützen, Rechtfertigungen für ihre Entscheidungen zu konstruieren.

Insgesamt verweisen beide Experimente auf die wichtige Funktion, die sowohl interne als auch externe Präferenzen und Normen als Richtungsziele für den Entscheidungsprozess haben. Die Bedeutung dieser Befunde für den Einsatz von Rechtfertigungsdruck als Kontrollmittel in Organisationen wird diskutiert.

SUMMARY

The thesis investigated the effects of accountability on information search, evaluation and integration in multiattribute decision making, combining classical social psychological with decision-analytical approaches. Accountability was conceptualised as a motivating agent, that is, as a variable that introduces particular goals into decision makers' information processing. The exact nature of these goals was argued to depend on the particular type of accountability that is created. Accountability demands were suggested to consist of two core components, (1) a feeling of personal responsibility for potentially negative consequences of the individual's behaviour, and (2) evaluation apprehension. However, only when the individual is made accountable to an external audience, are both components present. Under internal accountability, that is, a need to be able to justify one's behaviour to oneself, evaluation apprehension is largely absent, due to the absence of an external audience.

A new process model of accountability (PMA) was introduced whose main premise is that accountability can make the decision process both more complex and more biased and therefore has not the exclusively beneficial effects that it is commonly afforded. The PMA proposes that the personal responsibility cues inherent in accountability demands make accuracy goals salient, which the decision maker tries to meet by increasing the depth of information search and complexity of information integration. Evaluation apprehension cues, on the other hand, are assumed to make directional goals salient, that is, goals that drive the decision process towards a particular conclusion, for example, towards supporting an alternative that is assumed to be preferred by the external audience individuals are accountable to. A directional goal may also be introduced by a previous commitment to an alternative; in this case the decision process is biased towards the previously chosen alternative.

Some predictions derived from the PMA were tested in two empirical studies, both of which employed a similar paradigm. Participants searched information twice, once before they made a preliminary and once before they made a final decision. The information contained in the second search set was largely redundant with that of the first set, in order to be able to test whether information search after a previous

commitment would be biased to support the chosen alternative. Unidimensional value and weight judgements were elicited several times during the experiment, in order to test whether participants would change them in such a way as to increase the evaluative advantage of their chosen alternative over any non-chosen alternatives. The latter is predicted by Svenson's (1992, 1996) differentiation and consolidation theory, which has applied Festinger's (1957, 1964) theory of cognitive dissonance to multiattribute decisions.

The first study investigated the effects of pre- and post-decisional internal and external accountability on the decision processes observed during a fictitious personnel selection decision. No external norms or preferences were made explicit. As predicted, both internal and external accountability tended to result in a deeper and more complex information search prior to the first decision. Pre-decisional internal accountability was found to keep information processing after a preliminary decision relatively balanced. However, when internal accountability was only introduced after a previous commitment had been made, participants' information search and evaluation was more strongly biased towards supporting the previously chosen alternative than that of externally accountable participants. The analysis of the extent of bias in participants' information evaluation yielded rather weak effects, which suggested that under external accountability, individuals tended to decrease the evaluative difference between their chosen and average non-chosen alternative whenever their evaluation was obvious to others, for example, in holistic suitability ratings of alternatives, but tended to increase this difference when their evaluation was less obvious, for example, when unidimensional value and weight judgements were combined in a linear additive MAUT (Multiattribute Utility Theory) model prediction of the overall evaluation of alternatives. This suggested that externally accountable participants tried to avoid a strong public commitment to a particular alternative when the reactions of their external audience to this commitment could not be anticipated. Hence, impression management concerns seem to play an important role in determining the responses of individuals' who have been made externally accountable.

This latter conclusion was also supported by a second study, which investigated pre-and post-decisional accountability to different interest groups. Participants were asked to decide which software engineering tool they would acquire for a company, and were either made accountable to the future users of the tool or the management of the company. As predicted by the PMA, an adjustment to the preferences of the external audience was only observed when accountability had already been introduced before, but not when it had only been introduced after a preliminary decision. This adjustment was primarily observed in overt, and to a lesser extent in covert, judgements. When participants' preferences conflicted with those of the audience they had been made accountable to, however, both overt and covert judgements showed signs of bolstering, where participants increased the evaluative difference between their chosen and average non-chosen alternative and searched information that supported their chosen alternative over their non-chosen alternatives. At the same time, both pre-and post-decisional accountability resulted in deeper and more complex information processing compared to no accountability, suggesting that accountability will indeed increase both depth and directional bias in individuals' information processing. The increased depth of processing may help individuals to construct sophisticated justifications for their choice.

Taken together, both studies demonstrated the special role of internal as well as external preferences and norms in directing the individual's decision process. The implications of these findings for the use of accountability as a means of control in organisations were discussed.

CONTENTS

Figure captions

Table headings

List of equations

CHAPTER 1: INTRODUCTION

”Often anticipations of acceptable justifications will control conduct. (‘If I did this, what could I say? What would they say?’) Decisions may be, wholly or in part, delimited by answers to such queries.”

(C. W. Mills, 1940, p. 907)

Decisions are considered to be expressions of a decision maker’s value system, that is, to reflect the decision maker’s preferences. These preferences, however, are not stable but strongly context-dependent. Decision research has identified a variety of context variables which strongly influence both the process of arriving at a decision and the resulting decision (Payne, Bettman & Johnson, 1993). These are, for example, variables of the task context, such as the complexity of a decision problem or whether or not a decision has to be made under time pressure. As the above statement by Mills suggests, variables of the *social context* in which a decision is made must be assumed to be at least equally influential. As Tetlock (1992) points out, decisions are not usually made in a social vacuum. Often decisions affect not only the person who makes them but also other people, and those others expect that the consequences a decision will have for them are taken into account when it is made. Also, important decisions are often the product of the decision maker interacting with other people, asking them for and listening to their opinions, or they are made by a group altogether. Whereas social psychology has recognised the impact of the social environment on individual behaviour and has made it the focus of its study, this impact has largely been ignored by cognitive and economic approaches to decision-making. It is a goal of this thesis to integrate social psychological and cognitive approaches to decision-making, in order to achieve a more complete understanding of decision-making in a social environment.

The work presented here focuses on accountability as an important social context variable. The concept of accountability, which deals with the demand to justify one’s

behaviour, can be regarded as one of the oldest concepts in science (Schlenker, Weigold & Doherty, 1991), yet it has not received much empirical attention until recently. Both the extent to which a person's behaviour is affected by being accountable for it and the nature of these effects including their mediating processes has long been unknown. These issues will be addressed, with a focus on decisions as important constituents of human behaviour. In particular, the aim is to find an answer to the following questions: How does accountability affect the process of arriving at a decision and its outcome? How are information search, evaluation and integration influenced by the knowledge that a decision has to be justified to oneself and/or other people?

In recent years, accountability has come to be regarded as a remedy against inefficiency and ineffectiveness and even as a means to overcome the general crisis of modern society (Bayertz, 1996). Proposals for improving the efficiency and effectiveness of social services, for example, usually advocate increasing accountability (Newman & Turem, 1974; Ross & Joelson, 1986; O'Neill, 1989). This is the case although the effects of accountability on behaviour in general and decision-making processes in particular are not yet fully understood and there are good reasons to believe that, under certain circumstances, holding people accountable may have effects contrary to those intended. Adelberg and Batson (1978), for example, were able to show that, in the context of a financial resource allocation, when applicants' needs exceeded available resources, making the decision-maker accountable to either the provider or the recipient of resources led to a less effective use of resources than did no accountability. It is essential, therefore, to know more about the effects of accountability on decision-making and potential moderators of these effects before any such advice can be given. It is the aim of this thesis to contribute to the understanding of these processes.

The approach to decision making taken here is a descriptive one, that is, in contrast to normative theories of decision making which assume that context variables do not and should not affect the decision process or its outcome, the significance of such variables is recognised and the effects of accountability as one of these variables investigated. Descriptive theories of decision making often stress the importance of a decision maker's *goals* or *motives* for how decisions are made. The notion that motives affect reasoning processes is an old one in social psychology (Lewin, 1938; Festinger, 1957; Heider, 1958). General motives that have been proposed to influence reasoning processes include avoiding inconsistency in the cognitive

system (Festinger, 1957; Aronson, 1968), protecting or enhancing one's self-esteem (Steele, 1988), creating a favourable impression on others (Schlenker, 1980; Tedeschi, 1981; Baumeister, 1982) and avoiding stress and anxiety (Janis & Mann, 1977). These motives may be made salient by particular features of the situation and create specific processing goals. Accountability may be regarded as one such feature.

Abelson (1963) distinguished between hot and cold cognition. Hot cognition describes affect-laden, motivated information processing, whereas cold cognition refers to the normal, non-motivated operation of the human information processing system. Although very plausible, the idea of motivated, hot cognition has been controversial in the past. In the mid 1970s and early 1980s many effects previously accounted for in terms of motivation were explained in non-motivational, cognitive terms. It was suggested that motivated reasoning phenomena could be entirely explained by the nature of human information processing mechanisms and that it was not necessary to invoke any motivational explanations (Miller & Ross, 1975; Ross, 1977; Nisbett & Ross, 1980). This cold cognition approach assumed that people draw certain conclusions not because they are motivated to do so, but because these conclusions seem more plausible, given certain prior beliefs and expectancies (Kunda, 1990). An example for a purely cognitive mechanism that may underlie biased judgements and decisions is that people can only process information that is available to them. If, for some reason, the available information is insufficient or not representative, a biased conclusion will be drawn.

The hot versus cold cognition debate is an on-going one since it is often very difficult to determine whether a certain effect is due to a motivational or a purely cognitive mechanism. Abelson and Levi (1985) give the example of a theory-driven, biased information search which may reflect a motivational bias in that a decision maker may want to confirm his or her initial hypothesis but could also reflect a cognitive bias in that the decision maker's initial hypothesis may direct his or her attention to supporting information. Authors such as Tetlock and Levi (1982) have therefore argued that the hot versus cold cognition controversy cannot be solved, at least not in certain domains.

Nevertheless, the evidence that certain motives or goals can influence information processing seems to be indisputable. There have been efforts recently to put forward a

compromise view, for example by Kruglanski (1996), who suggests to get rid of the distinction between motivated and non-motivated cognition altogether and adopt a synergistic perspective. According to Kruglanski, "motivation and cognition are inextricably intertwined in that nearly all motivation encompasses cognitive aspects and nearly all cognition encompasses motivational aspects" (1996, p. 493). He proposes that, instead of asking whether motivation exerts any effects on cognition, one should concentrate on answering a slightly different question, namely what it is about motivation that helps us understand cognition and vice versa. Kunda (1990) proposes that motivation may affect reasoning through reliance on a set of cognitive strategies from which the individual chooses the one that best suits his or her goals and which provide the mechanisms through which motivation can affect reasoning. This is a view of the individual as a motivated tactician, who regulates his or her own cognitive processes in order to fulfil certain needs and achieve certain goals (Fiske & Taylor, 1991; Showers & Cantor, 1985; Pyszczynski & Greenberg, 1987; Kruglanski & Meinholdt, 1990; Baumeister & Newman, 1994). Baumeister and Newman's (1994) self-regulation model will serve as a basis for a theoretical framework that describes how the individual may direct his or her decision processes as a consequence of accountability pressures and will be presented in some detail in Chapter 4.

Although the work presented here investigates decision processes with a descriptive rather than normative focus, the methodology employed is largely based on a normative decision theory, multiattribute utility theory (MAUT, Edwards, 1971, 1977; Raiffa, 1969; Keeney & Raiffa, 1976; von Winterfeldt and Edwards, 1986). We will therefore continue with an outline of what decision making entails and introduce MAUT as both a theory of how decisions should be made and a set of methods to guide the individual in making a rational decision. We will then continue with a discussion of accountability and its significance in various real-world contexts before concluding this chapter with a sketch of the chapters to come.

Decision processes

A decision is commonly conceived of as choosing one alternative out of a number of available alternatives. This is assumed to be an intentional act, preceded by a period of conscious reflection about the different alternatives. There is no general agreement among decision researchers about how people actually make decisions (Huber, 1989), but it may be argued that there are three fundamental processes involved: *Information search*, *information evaluation*, and *information integration*. When making a decision, the individual first has to find out about various aspects or consequences of the available alternatives. This is the process of information search. The next important step is to evaluate these consequences in terms of how well they serve one's goals, and, finally, to integrate these evaluations to arrive at a decision.

Not all decisions, however, are made in this way. Decisions can vary considerably with regard to the extent of conscious cognitive effort that is expended when making them. Often we are not even aware of making a decision. This is the case when decisions have become largely habitual and automatised, because they have to be made often and are not very important. The decision maker simply chooses what he or she has chosen before and what turned out to be successful (so-called level 1 decisions, Svenson, 1990). An example for this type of decision would be choosing which brand of coffee to buy in a supermarket.

Only when decisions are perceived to be important and difficult does the decision maker carefully consider all available alternatives and make trade-offs between the various advantages and disadvantages of alternatives (Borcherding, Schmeer & Weber, 1995). Whether a decision is perceived as important, is mainly a function of its consequences (Borcherding, 1991). If the consequences are serious, affect many people and will be present for a long time or even irreversible, the decision is considered to be *important*. Decisions are perceived to be *difficult* when the decision maker has multiple conflicting goals and no alternative is clearly superior to the others in terms of achieving these goals. This is usually the case for *multiattribute* decisions. Multiattribute decisions are decisions between alternatives, each of which has several consequences. These consequences occur in different dimensions (attributes) and are called outcomes. An example for a multiattribute decision is a

decision about which house to buy. The houses on offer may differ with regard to price, size, state of repair, distance to work and distance to shopping facilities; they have different outcomes on these attributes. In such a situation, the decision maker usually has multiple objectives: He or she wants to buy a house that is at the same time inexpensive, has an optimal size, is in a good state of repair, close to work and close to shopping facilities. Most likely there will not be an alternative that serves all these goals best. For this reason, multiattribute decisions are also called decisions with conflicting objectives (Bell, Keeney, & Raiffa, 1977). In order to arrive at a decision, the decision maker has to compare all aspects and trade them off against each other. The resulting trade-offs are likely to differ between individual decision makers and, therefore, are open to criticism. Hence, multiattribute decisions are typically decisions we are accountable for and therefore are the type of decision investigated here. People do not normally feel accountable or are made accountable for unimportant and easy decisions, unless they lead to unexpected negative consequences.

Multiattribute utility theory

Multiattribute utility theory (Edwards, 1971, 1977; Raiffa, 1969; Keeney & Raiffa, 1976; von Winterfeldt and Edwards, 1986) is a prescriptive theory that sets out how multiattribute decisions *should* be made (in contrast to descriptive theories of decision making which deal with how decisions *are* made). MAUT can therefore also serve as a support tool for individual or group decision making. As a prescriptive theory MAUT takes into account the subjective nature of a decision maker's goals and attempts to identify the choice that serves these subjective goals best, in contrast to normative decision theories which propose general rational principles that all decision makers should follow. MAUT recognises that multiattribute decision situations typically make heavy demands on the decision maker's information processing capacities (Borcherding, 1983, Hogarth, 1987) and suggests simplifying the decision process by decomposing the decision problem into its constituting components, evaluating these components separately, and then recomposing them to arrive at a decision (Keeney & Raiffa, 1976). Instead of having to cope with an *n*-dimensional decision problem, decomposition of the problem allows the decision maker to deal with *n* one-dimensional problems separately.

The application of MAUT usually involves four major steps whose order, apart from the first and the final step, is arbitrary: (1) structuring objectives, (2) eliciting preferences, (3) scoring alternatives, and (4) aggregating preferences to arrive at an overall evaluation for each alternative. The main goal of structuring objectives is to select dimensions on which the alternatives are to be evaluated. The decision maker specifies his or her objectives (e.g., buying a house which is inexpensive, of optimal size, in a good state of repair, close to work, and close to shopping facilities), which are then build into a value tree, with the overall, general objective (e.g., buying a house) at the top and increasingly specific objectives (e.g., optimal size) at lower levels of the tree. The objectives at the bottom level of the tree have to be measurable and are usually referred to as attributes (see Keeney & Raiffa, 1976, for desirable properties of a set of attributes).

Once a value tree has been created the decision maker's preferences are measured. Preferences include the relative importance of attributes, referred to as attribute weights, and the value or utility of certain outcomes on the specified attributes. The term 'value' is normally used when dealing with decisions which have more or less known consequences (e.g., which house to buy), whereas 'utility' is reserved for decisions involving uncertainty to such an extent that it warrants representation in the decision analysis (e.g., which shares to buy). When values or utilities are measured, the decision maker specifies the relationship between particular outcomes and the extent to which these outcomes satisfy the objective expressed by their corresponding attribute. For example, the decision maker may be asked to specify, for a range of sizes, how good (or bad) they are in terms of the wished-for size. The resulting value or utility functions allow comparing aspects of alternatives across different dimensions, that is, they establish commensurability across different attributes. Values or utilities are based on the same scale, unlike measurements made on an attribute level which usually differ in the units employed (e.g., size may be expressed in m^2 and distance to work in km) and therefore make trade-offs between attributes difficult.

There are a number of different ways in which values and utilities may be measured (see von Winterfeldt & Edwards, 1986). The simplest one, direct rating, involves the decision maker rating the outcomes of alternatives on an attractiveness scale which is usually anchored by the worst and best plausible outcome. In the house-buying example, the decision maker may be asked how good he or she perceives a distance to work of 10 km to be, when the best

outcome would be living 0.5 km away and receives a score of 100, and the worst outcome would be living 50km away and receives a score of 0. Alternatively, the scale may just have verbal anchors, for example, ‘very bad’ and ‘very good’.

Attribute weights describe the importance of each attribute *relative* to the other attributes and the decision maker’s overall objective. The term ‘relative’ is important here; attribute weights specify the contribution of each attribute to the overall evaluation of an alternative in the given decision situation and not some abstract, general perception of importance[1]. Weights, like values or utilities, can be elicited in different ways. One method, the swing method (von Winterfeldt and Edwards, 1986) will be described in more detail in Chapter 6. For a summary of weight elicitation methods and a discussion of their respective merits and disadvantages, see Borcherding, Eppel, and von Winterfeldt (1991) and Borcherding, Schmeer, and Weber (1995).

MAUT's third step, the scoring of alternatives, involves collecting information about the outcome of each alternative with respect to each attribute. For example, the decision maker who wants to buy a house would have to gather information about the price, size, state of repair, distance to work and distance to shopping facilities of each house he or she considers. This, together with the elicited value or utility functions, allows determining the perceived value or utility of each alternative with regard to each attribute and is an important prerequisite for the final step, the aggregation of preferences to arrive at an overall evaluation for each alternative.

Different aggregation (or decision) rules may be applied. By far the most common and the simplest aggregation rule is the *weighted additive rule*, whereby the overall evaluation of an alternative is a function of the value or utility of an outcome multiplied by the weight of the corresponding attribute, added up over all attributes. Other aggregation rules include the multiplicative and the multilinear model (see Borcherding, 1983, for a detailed discussion of

1 This implies that attribute weights should change with the range of outcomes of the considered alternatives. If the range of outcomes is small (e.g., the price of houses on offer only varies by DM 500), the price attribute should receive a small weight. If the range of outcomes is large, on the other hand (e.g., DM 50,000), the weight of the price attribute should increase. Decision makers, however, often show range insensitivity, that is, they do not make sufficient adjustments when outcome ranges change.

these models). All models have in common that they involve calculating some sort of weighted average which reflects the expected value (EV) or subjectively expected utility (SEU) of an alternative. A formal representation of the weighted additive rule is shown in Equation 1.1:

$$V(A_i) = \sum w_j \cdot v(x_{ij}) \qquad \text{(Eq. 1.1)}$$

where V(Ai) is the overall evaluation of an alternative i, wj is the weight for the jth attribute, xij is the outcome for alternative i on attribute j, and v(xij) is the value or utility of alternative i's outcome on attribute j.

The aggregation rule is used to calculate an overall evaluation for each considered alternative. The alternative that receives the highest overall evaluation maximises EV or SEU and should be chosen by a rational decision maker.

Compensatory vs. non-compensatory decision rules

Decision rules are typically categorised as compensatory or non-compensatory. A *compensatory* decision rule allows a bad value on one attribute to be compensated for by a good value on another attribute, that is, it permits trade-offs. MAUT's aggregation rules, such as the weighted additive rule, are examples of compensatory decision rules. This is not surprising, as MAUT explicitly deals with how trade-offs between attributes should be made. MAUT, as a prescriptive theory, is concerned with making rational choices, and rational choices are assumed to be those (and only those) which maximise expected value or utility. Value or utility maximisation, by nature, involves a compensatory approach; all outcomes are considered and weighed by the extent to which the decision maker is willing to trade them off against each other (in the case of value maximisation, Hogarth, 1987) or by the probability with which they are obtained (in the case of utility maximisation).

Although compensatory decision rules are best from a normative point of view, descriptive analyses of decision processes have repeatedly shown that decision makers do not use them very often. Decision makers prefer non-compensatory decision rules, especially when the decision to be made is complex (Payne, Bettman, & Johnson, 1993) or not highly involving (Kahn & Baron, 1995). As the term suggests, *non-compensatory* decision rules do

not allow compensation or trade-offs between attributes. Examples for non-compensatory decision rules are the conjunctive, the lexicographic and the elimination-by-aspects rule (see Svenson, 1979, for others). The conjunctive rule involves the setting of cut-off points for each attribute; only alternatives whose outcomes exceed the cut-off point for each attribute may be chosen. The lexicographic rule prescribes the choice of the alternative with the best outcome on the most important attribute. If alternatives have equally attractive outcomes on the most important attribute, they will be compared on the attribute that is the next most important until only one alternative remains. The elimination-by-aspects rule, suggested by Tversky (1972), combines aspects of both the conjunctive and the lexicographic rule. Similarly to the lexicographic rule, alternatives are compared on attributes in the order of attribute importance, but, similarly to the conjunctive rule, only alternatives whose outcomes exceed certain cut-off points survive the elimination process.

Why are decision makers more likely to use non-compensatory than compensatory decision rules? A major reason for the attractiveness of non-compensatory rules is that they are usually much less demanding in terms of the cognitive effort they require. Non-compensatory decision rules involve fewer and less complex comparisons than compensatory decision rules. Also, the decision maker does not have to establish commensurability across attributes in order to be able to apply them. This advantage becomes particularly important when information load is high due a large number of alternatives and/or attributes, and the decision maker's information processing capacities are stretched (Payne, Bettman, & Johnson, 1993). Simon (1955) introduced the concept of bounded rationality which recognises the decision maker's wish to make a rational decision but at the same time keep cognitive effort minimal. Simon suggests that a decision maker often engages in a process of satisficing rather than maximising or optimising, that is, he or she looks for a satisfactory alternative rather than the best one. This process involves the application of a heuristic very similar to the conjunctive rule; the first alternative that meets certain criteria or so-called aspiration levels is chosen. Satisficing and other non-compensatory decision rules have the benefit of reducing computational demands, but they also carry a cost; they may result in the rejection of an alternative that is superior to the chosen one. A satisficing decision maker is willing to accept this risk in exchange for the benefits of the chosen rule. This idea of Simon's greatly influenced descriptive theories of decision making and resulted in the proposal of

various cost-benefit frameworks in the decision literature, such as Beach and Mitchell's (1978) strategy selection model and Payne, Bettman, and Johnson's (1993) effort-accuracy framework. A similar model from the social cognition literature is the cognitive miser model (Taylor, 1981, Taylor & Fiske, 1981) which suggests that, because people have limited information processing capacities, they take shortcuts whenever possible.

Hogarth (1987) points out that, apart from allowing the decision maker to avoid the cognitive effort of trade-offs, the use of non-compensatory decision rules also allows him or her to avoid the emotional conflict that typically accompanies trade-offs. Trade-offs force the decision maker to consider questions he or she would like to avoid, for example, how big a cut in salary would I be willing to accept for working 10 hours less per week. People show a particular aversion to trade-offs that involve human lives, such as decisions about how many lives saved are worth a certain investment into road safety campaigns or into safety measures for a nuclear waste disposal site (Lichtenstein, Gregory, Slovic, & Wagenaar, 1990). The application of non-compensatory decision rules is an attempt to reduce the conflict by avoiding trade-offs altogether, whereas the application of compensatory decision rules shows a conflict-confronting strategy.

Unless process tracing methods are applied which allow the researcher to monitor how decisions are made, it is difficult to infer which rule a decision maker used to arrive at his or her choice. The fact that a certain rule predicts a choice well, however, does not necessarily mean that this rule was really applied. Linear models, such as the weighted additive model, usually are good predictors of choices, even if the decision rule used was not a compensatory one (Dawes & Corrigan, 1974; Dawes, 1979).

Having outlined general principles of decision theory, we will now return to our original interest, namely accountability as a variable that affects decision processes.

Accountability as a norm enforcement mechanism

According to Schlenker, Weigold and Doherty (1991), accountability "refers to being answerable to audiences, oneself included, for performing up to prescribed standards" (p. 110). This definition incorporates two important aspects of accountability: (1) accountability

comprises a *need to explain one's behaviour* either to oneself or to an external audience, and (2) the explanations brought forward are oriented at and will be evaluated with respect to *certain prescribed standards or norms*. The latter aspect implies that accountability can be understood as a norm enforcement mechanism. It is considered to be crucial for social control and self-regulation (Schlenker, Weigold & Doherty, 1991). According to Tetlock (1985a), accountability provides the social psychological link between individual decision makers on the one hand and the social system to which they belong on the other. For a society it is vital to solve the problem of how to control and co-ordinate the actions of individuals who are capable of monitoring and controlling their own actions and who have different interests and goals (Semin & Manstead, 1983; Tetlock, 1985a; Schlenker, Weigold & Doherty, 1991). Accountability causes people to expect that others will judge their behaviour with regard to prevalent norms of conduct and that rewards and punishments will be distributed accordingly. This puts a certain constraint on individual behaviour and allows people to exert legitimised control over each other's behaviour (Schlenker, Weigold, & Doherty, 1991). In this context, it is also important to note the significance of accountability for maintaining good relationships with the social environment (Buttny, 1993). Accounts help to repair 'fractured' social interaction, that is, interaction which has been disrupted by the fact that social norms have been violated or are believed to have been violated (Semin & Manstead, 1983). When accounting for their behaviour, individuals may try to convince the other party that they could not have foreseen the consequences of their behaviour and therefore are not responsible for them or that these consequences are not as negative as they appear to be. Otherwise, people may apologise for their behaviour. When the other party accepts the accounts or apologies brought forward, normal interaction can be resumed. Accounts also allow individuals to negotiate rules and norms (Buttny, 1993). Although the social control function of accountability constrains the individual from deviating from societal norms of conduct, the question of how and whether these norms apply to the individual can be negotiated in social interaction, especially through accounts.

Accountability not only provides a mechanism for social control, but also for *self-regulation*. Even when not explicitly prompted by other people to do so, individuals take into account how their behaviour may be evaluated in the light of certain standards and norms. They consider how defensible certain actions are and adjust their behaviour accordingly, that

is, they organise their behaviour in such a way that it can be accounted for to others (Semin & Manstead, 1983). Individuals also monitor their behaviour, evaluate the resulting performance and administer self-reward or self-punishment, thereby exerting self-regulation (Bandura, 1977).

To illustrate the high significance of accountability for various domains of public life, we will now discuss four contexts in which accountability features particularly strongly: the law, politics, organisations, and professions.

Accountability contexts

Legal accountability

Given the control function of accountability and its central role for the functioning of a society, it is not surprising that it has a particular significance in legal contexts. Very often accountability also means legal responsibility. Legal systems rest on the assumption that people are autonomous actors who are able to control their actions and account for them. When an offence has been committed and a person is brought before a court of law, a first step in order to establish whether a defendant is guilty and therefore liable to punishment, is to call him or her to account, that is, to demand a more detailed explanation from him or her about what happened and to give reasons for his or her behaviour (Ross, 1975). Only when responsibility of a person for his or her deed has been established, following certain rules specified in the law, the person can be found guilty and sentencing can take place. This implies that accountability in a legal context not only entails accountability of the defendant to a court of law, but also accountability of the court of law to the defendant. The court of law has to give reasons for arriving at a certain verdict. Thereby a reciprocal relation or a contract is established between the individual and the legal system: the individual is accountable for the consequences of his or her actions, but only if the law is not subject to arbitrary change or enforcement (McKeon, 1957). Legal trials can be understood as a social institutionalisation of the everyday process of accusation and justification. This form of social institutionalisation has in turn exerted considerable influence on the procedure of accounting for oneself in an

everyday context. The process of accounting for one's behaviour often resembles a legal trial (Bayertz, 1996).

Political accountability

Clearly, accountability is also an integral part of politics. Accountability can be regarded as a central feature of a democratic political system. Members of parliament and government are accountable to their constituencies and control institutions. The idea of an accountable and responsible government has developed and has been applied to an increasing extent since the 19th century. A responsible government „operates within a framework of law in which official action and control are reasonably predictable" and „reflects the attitude of its people through institutions which provide for the regular election of personnel and regular review of policy" (McKeon, 1957, pp. 23-24).

Accountability in a political context results from the fact that someone is a representative of or agent for the general public and being given a mandate to act in the interests of society. Elections provide the means for the individual to administer rewards or punishments for politicians' performance at this task. Politicians who have not been able to fulfil their task in an acceptable way will usually not be elected again. They will therefore try to convince the public that they have acted or will act to the benefit of the society.

According to Goldmann (1971), politics to a very large extent consists of justifying past and future decisions, and of criticising justifications brought forward by others. That the necessity of defending their decisions is a key constraint on how policy-makers choose among available alternatives, has been demonstrated in a number of case studies of important foreign politics decisions (Anderson, 1981). After Egypt nationalised the Suez Canal in 1956, „the British government consciously tried to create a situation in which the use of force against Egypt would be acceptable in terms of certain international norms" (Goldmann, 1971, p. 44). British leaders tried to hide their collusion with Israel and delayed the invasion of Egypt until after the initial Israeli strike. Another example is President Kennedy's demand that the 1961 Bay of Pigs invasion of Cuba should not be traceable to U.S. armed forces, because he had publicly pledged not to violate the norm of military non-intervention. This had important consequences in that the invading forces had a very inefficient air force and that the whole

operation failed (Goldmann, 1971, Janis, 1982, Anderson, 1981, Tetlock, 1985a, 1992). As an example from domestic politics, one may cite the attempt of German politicians across all parties to justify a law waiting to be passed by the German parliament in June 1995 that intended to couple the allowances of members of parliament (MPs) with the income of high court judges. As this would have meant a substantial rise in income for MPs at a time when the majority of the population suffered losses in their real income, MPs were given 'justification support' (Argumentationshilfen), a brochure they were supposed to present to their home constituencies that presented arguments in favour of the planned pay-rise. Owing to continuing public opposition, however, the law did not attract the necessary majority of votes.

Tetlock (1985a, 1992) put forward the idea that all human beings act like politicians insofar as they make judgements and decisions depending on whether they are justifiable to significant others. In his social contingency model of judgement and choice he uses the metaphor of the individual as a „politician whose primary goal is to maintain the positive regard of important constituencies to whom he or she feels accountable“ (1992, p. 332).

Organisational accountability

Accountability demands in organisations occur on different levels and to different extents, depending on the structure of the organisation and the position a person occupies in it. The more structured, particularly, the more simply and hierarchically structured an organisation is, the more likely is it that accountability demands are salient. In a simply structured organisation top level managers have the final say over all important decisions. What Mintzberg (1989) calls the ‘operating core’, namely the manual workers and people who provide the services, are directly accountable to the top level managers. Here accountability arises from the superior-subordinate relationship. Romzek and Dubnick (1987) call this kind of accountability *bureaucratic accountability*. As an organisation gets larger, the distance between top level managers and their subordinates increases and makes it more difficult to ensure accountability of subordinates to organisational leaders. Rules and regulations typically replace direct supervision as a means to hold people accountable (Jos & Tompkins, 1994). The more unstructured, diversified, or divisionalised an organisation is, the more difficult it becomes for central administrators to exert control. Also, for subordinates it

becomes increasingly difficult to identify whom they are accountable to. The fact that subordinates may be accountable to many people with perhaps conflicting goals and expectations not only makes it difficult for them to decide which expectations to follow, but also often makes it more likely that they will be able to protect themselves from accountability demands altogether. This is a problem especially prominent in public administration. Public administrators often find themselves in a situation where a fundamental question is to whom they should be answerable and for what (Jos & Tompkins, 1994). Potential constituencies include "the general public, elected officials, agency heads, agency clientele, other special interest groups, and future generations" (Romzek & Dubnick, 1987, p. 229). The constituencies in turn may lack the resources and cohesiveness to consistently monitor administrative implementation. We encounter here what Eisenhardt (1989) has called the fundamental agency problem, „the desires or goals of the principal and the agent conflict and it is difficult or expensive for the principal to verify what the agent is actually doing" (p.58).

Top level managers not only make accountability demands on their subordinates but are in turn accountable to external groups, like share-holders or other providers of resources (Staw, 1980). Thus, according to Jos and Tompkins (1994), organisational leaders are in a position where they both make others in the organisation accountable and themselves are made accountable by others outside the organisation. Pfeffer (1981) also stresses the importance of justifiability in managerial decision-making. He sees the primary task of managers to „provide explanations, rationalizations and legitimation for the activities undertaken in the organization" (p.4), that is, to make what is going on in the organisation meaningful and acceptable to its members and to the world outside. In his view, only a manager who is able to convince both internal and external observers that the organisational activities are consistent with prevailing social norms and values can be successful. According to this view, managers in the same way as politicians search for courses of action that can easily be justified to important constituencies.

The need to justify their actions may also result in organisations developing 'justificatory myths' (Meyer & Rowan, 1977), that is, justifications that are intended to convince a sceptical public. This is particularly likely when the organisation is under high pressure from external groups.

Professional accountability

Jos and Tompkins (1994) speak of professional accountability when a member of a certain profession is either accountable to a formal professional body that has developed its own standards and codes of conduct, or when there exists an internalised set of norms and principles that reflect an individual's 'professional conscience' which he or she tries to follow. An example that comes readily to mind here is the medical profession. More generally, professionals whose judgements or actions may lead to severe negative consequences for people other than themselves are usually faced with the prospect of being held accountable. Apart from the medical profession, this is true for lawyers, accountants and auditors, for example (Messier & Quilliam, 1992).

Accountability is an important aspect of the auditor's judgement and decision environment (Gibbins & Emby, 1985). This is recognised not only by people outside the profession but also by auditors themselves. When practising auditors judged the importance of various factors in exercising good professional judgement, the quality viewed most important was that the decision could be justified (Gibbins & Emby, 1985, Emby & Gibbins, 1988). Auditors usually make decisions in the knowledge that their decisions will be reviewed by partners in the firm and that explanations for their decisions will be called for (Johnson & Kaplan, 1991). The review process is considered to be an integral part of a firm's quality control procedures. Pressure on auditors to justify their judgements and decisions, however, might also be exerted by clients (Lord, 1992).

Professional accountability arises from a person having special expertise. A control system that stresses professional accountability is especially appropriate when the task at hand is complex and technically difficult (Romzek & Dubnick, 1987). Under such circumstances organisations must rely on skilled and expert employees who expect to be held fully accountable for their actions but in turn expect that they are given sufficient discretion to fulfil their duties. This means that a good deal of control over organisational activities is placed in the hands of expert employees. Romzek and Dubnick (1987) argue that the Challenger accident in 1986, when the space shuttle Challenger exploded shortly after launch and seven crew members died, would very likely not have happened had NASA relied on a system of professional accountability rather than bureaucratic and political accountability.

The decision to launch the shuttle despite adverse weather conditions was made under political pressure to achieve planned flight rates. This pressure had been exerted by the White House, the Congress, military and private sector customers, and repeated media reports of shuttle delays. The management of Morton Thiokol, NASA's main shuttle program contractor, signalled their consent to the decision to launch despite the reservation of their engineers, because NASA was their main customer and in the process of reviewing its contracts with the firm. At NASA, lower-level supervisors were reluctant to report problems voiced by their engineers upward, because they did not want to be perceived as the ones who endangered the success of the program. The final decision to launch was made by managers who did not have the necessary information and expertise to make a correct decision. If a professional accountability system had been given more significance in the decision-making process, as would have been appropriate given the technically difficult task NASA faced, the recommendation of the engineers not to launch would very likely have been followed and the disastrous accident avoided (Romzek & Dubnick, 1987).

Accountability as a motivating agent

The above discussion of the different societal contexts in which accountability plays an important role should have given the reader an idea of its significance for the analysis of individual behaviour in a social environment. Accountability is a universal feature of behavioural environments (Tetlock, 1985a, 1992) and therefore relevant to a wide range of phenomena investigated in the social sciences. In social psychology these phenomena include, for example, conformity, social comparison, social exchange, performance in the presence of others, attribution, impression management, and identity development. Schlenker, Weigold and Doherty (1991) even go so far as to suggest that problems in dealing with accountability are at the core of most dysfunctional behaviours and that accountability therefore also bears upon the development of psychopathological symptoms. The theoretical analysis of the accountability construct, following in Chapter 2, will attempt to clarify the link between accountability and some of the aforementioned phenomena.

This thesis will argue that accountability is a motivating agent which typically makes two different needs salient, a need to be accurate and a need to arrive at an easily defensible

decision. Whether or to what extent one exceeds the other, depends on the particular situation the individual finds him- or herself in. It is therefore important to pay attention to the particular demands accountability manipulations create. It will also be suggested that the decision maker uses different cognitive mechanisms to try to achieve each of the goals made salient by accountability pressures, and that these mechanisms operate at different stages of the decision process. In particular, it is suggested that the decision maker will try to maximise accuracy by conducting an extensive information search and integrating it in a complex manner while he or she will try to arrive at an easily defensible decision by biasing information evaluation and integration in such a way that the alternative that the decision maker assumes appeals most to the people he or she is accountable to will be supported. Finally, while a motivation to be accurate and a motivation to make a decision that can be easily justified are often compatible, under certain circumstances they may conflict with each other. The ultimate effects of accountability on decision processes will depend on whether these goals conflict and if so, which goal will dominate. The investigation of accountability effects will combine a descriptive theoretical approach, primarily based on dissonance theory (Festinger, 1957, 1964) and differentiation and consolidation theory (Svenson, 1992, 1996) as the application of dissonance ideas to multiattribute decision situations, with a normative methodology as specified by MAUT.

Summary and overview

In this chapter we discussed the cognitive processes involved in decision making and introduced multiattribute utility theory (MAUT) as a theory which both sets out how decisions should be made and proposes a methodology to guide individuals in making rational decisions. We also discussed the main features of accountability and the important function it has for social control as well as self-regulation. Accountability was shown to be a pervasive aspect of our every-day with particular significance in the legal and political system and the work world. An initial proposal of how accountability exerts its effects on decision processes suggested that it is a motivating agent which makes certain information processing goals salient which, in turn, determine the cognitive strategies used by the decision maker to arrive at his or her decision.

In the next chapter, an analysis of the accountability construct will be presented, which discusses the different antecedent conditions of accountability and their link to other concepts in social psychology. Chapter 3 will review previous research on the effects of accountability on decision processes. In Chapter 4, we will introduce Baumeister and Newman's (1994) self-regulation model as a framework for the different mechanisms which underlie biased information processing as well as discuss the empirical evidence for these mechanisms, especially where multiattribute decisions are concerned. In Chapter 5, a process model of accountability will be presented which integrates empirical findings with some theoretical considerations from the previous chapters. Hypotheses derived from this model were tested in two empirical studies, which will be reported in Chapters 6 and 7. Finally, in Chapter 8 the presented findings will be summarised and implications for future research as well as possible applications discussed.

CHAPTER 2: THE ACCOUNTABILITY CONSTRUCT

In the decision making literature accountability has often been treated as a moderator variable, with relatively little interest in the concept itself and, with some notable exceptions, little awareness of how differences in the way accountability is manipulated may change the nature of the effect it has. The failure to recognise the complexity of the variable has impeded attempts to integrate empirical findings, which have suggested different and sometimes contrary effects of accountability on decision processes. Accountability, clearly, is not a single situational determinant that produces uniform effects on decision processes. Rather, for an individual to feel accountable, certain cues or antecedent conditions have to be present and each of these will have a certain effect on the way in which information is processed. A full understanding of the effects of accountability on decision processes must therefore be informed by an analysis of its antecedent conditions.

Antecedents of accountability

Semin and Manstead (1983) state that people will be asked to or spontaneously offer accounts for their behaviour only if a problematic act occurs, that is, if certain social rules or conventions are breached. In order for a situation to be recognised as such, several conditions have to be met (p. 36): (1) the rule(s) and conventions in question must belong to the repertoire available to the actor, (2) the rules and conventions must be taken seriously, (3) the actor must be aware of having broken the rule(s), (4) the rules must be recognised by the actor as being applicable to him- or herself, and (5) the actor must have some awareness that the breach in question has been witnessed by others.

Similarly, Schlenker (1986) suggests that accountability plays a major role whenever a predicament is encountered. He defines predicaments as "situations in which events have undesirable implications for the identity-relevant images actors have claimed or desire to claim in front of real or imagined audiences" (p.125). According to Schlenker, there are three conditions which normally accompany predicaments: (1) people believe their decisions or behaviours are publicly known or will become public knowledge, (2) people believe their

performance will be judged by others, and (3) people anticipate having to explain or justify their decisions or conduct to others.

Schlenker and co-workers (Schlenker & Weigold, 1989; Schlenker, Weigold & Doherty, 1991; Schlenker, Britt, Pennington, Murphy, & Doherty, 1994) have proposed a triangle model of accountability which specifies the conditions under which a perception of accountability develops and that determine its strength. According to the model, perceived accountability is a function of the combined strength of three elements and the connections or linkages between them: *prescriptions*, an *event*, and the actor's *identity*. These three elements are perceived as necessary to carry out the evaluation underlying accountability. Prescriptions (which can be either implicit or explicit) include information about what the individual should try to achieve (goals), in what way (rules), and what level of performance is satisfactory (standards). Perceived accountability is stronger, the more important the prescriptions are. Importance is perceived to be a function of both the extent to which the prescriptions are valued principles of conduct and any personal consequences of following them. The event is the individual's behaviour and its consequences. The more important the consequences, the stronger perceived accountability will be. Finally, the individual's identity is both a structure that contains ideas about what one is like and an active process of presenting oneself to be a particular kind of person (Schlenker, 1986). Perceived accountability will be stronger, the more strongly the situation bears on protecting or enhancing the individual's identity.

As mentioned above, perceived accountability does not only depend on the strength of the three elements but also on the strength of the linkages between them. The *prescription-event linkage* refers to the extent to which prescriptions for behaviour in the particular situation exist and are known. If prescriptions are clear and well-understood, the prescription-event linkage is strong, if they are "ambiguous, conflicting, obscure, and/or non-existent" (Sheer & Weigold, 1995, p. 596), the linkage is weak. The *prescription-identity linkage* describes the extent to which the prescriptions are binding for the actor. If the prescriptions clearly apply, for example, because the actor understands the consequences of his or her behaviour, occupies a certain role that implies adhering to them, or feels committed to them because of certain personal convictions, the prescription-identity linkage is strong. If, however, the prescriptions are inapplicable or the actor does not feel committed to them, the linkage is weak. Finally, the *identity-event linkage* refers to the extent to which the actor is

perceived to be personally responsible for the consequences of his or her behaviour, which, in turn, is a function of the intention to produce them and the control over them (we will discuss this last point in more detail below). The link is strong, if it is assumed that the actor had (or will have) both the intention to produce and control over the event, and weak otherwise.

Schlenker et al. (1994) provided support for the assumption that the perceived strength of accountability depends on the model's elements and the strength of their linkages. In a hypothetical scenario, actors were perceived to be most accountable when all linkages were strong, less accountable when one or two linkages were weak, and least accountable when all linkages were weak. When subjects were asked to gather information to determine whether the actor in a hypothetical scenario was accountable, 90% of the information requested was related to the proposed linkages of the model. Also, Sheer and Weigold (1995) found that when participants were asked to choose between accounts, they preferred those which weakened the linkages proposed by the triangle model of accountability.

In terms of factors which underlie perceived accountability, Semin and Manstead (1983) and Schlenker and co-workers seem to agree on the following, (1) negative consequences of an actor's behaviour (which may be intangible if they concern the actor's self and/or public image), (2) the actor feeling personally responsible for these consequences (or at least expecting to be made personally responsible), and (3) an anticipation of being evaluated. With regard to the first, negative consequences of behaviour, it should be noted, however, that people may be told explicitly and in advance of any action on their part that they will be accountable. Formal accountability procedures may be in place, especially in organisations and public administration, which require individuals to provide an account for their actions, regardless of the outcome. Antaki (1988) recognises that accounts are not only given to protect oneself in the event of a norm violation, but may also be descriptions, ordinary explanations, or self-reports of everyday activities. It may therefore be concluded that negative consequences are not a necessary condition of accountability. It seems reasonable to assume, however, that the consequences of the behaviour in question must be *potentially* negative and that the individual is aware of these potentially negative consequences. This implies the following conditions as antecedents of accountability:

(1) an actor's *personal responsibility* for *potentially negative consequences* of his or her behaviour (or an expectation to be held personally responsible), and

(2) *evaluation apprehension*.

We will continue with a discussion of personal responsibility and evaluation apprehension, as these variables have been investigated in some depth in other contexts and the results of these investigations make a valuable contribution to an understanding of the accountability concept.

Personal Responsibility

Accountability and responsibility are closely interrelated concepts. This is reflected in the fact that the terms ‘accountability’ and ‘responsibility’ are often used interchangeably. In languages other than English, for example, German, there is often only one term (Verantwortlichkeit) which is used for both. A person who is accountable for something is responsible for it. We have just argued, however, for the opposite; a person who is responsible for something is accountable for it (see also Greenfield, 1975). Thereby, responsibility is an antecedent condition of accountability and not identical with it.

Semin and Manstead (1983) refer to Ross (1975), who stresses the importance of making a clear distinction between the different meanings of responsibility. He suggests that responsibility can be (1) responsibility in the sense of being the person who can be “rightfully brought to account for” an act (p. 16), and (2) responsibility in the sense of being the person who can be “rightfully sentenced” for an act (p.20). The former is responsibility in the sense of having caused a particular outcome by one's behaviour and is closely related to accountability, the latter is responsibility in terms of legal liability. In some cases, people may be answerable not only for their own actions but also for the actions of others, by virtue of their role. For example, supervisors may be responsible for the actions of their subordinates (Graumann, 1994; Hamilton, 1978).

Personal causality and control as necessary antecedents of personal responsibility

Attribution research has investigated the conditions under which personal responsibility is ascribed to an actor. We will outline very briefly and selectively the conclusions that have been drawn (for a detailed discussion, see Fincham & Jaspars, 1980; Semin and Manstead, 1983; Shaver, 1985; Weiner, 1995a, b). The first contribution to an understanding of personal responsibility attributions was made by Heider (1958), who suggested five levels of responsibility for one's actions: (1) *global association*, whereby the actor is perceived to be responsible for outcomes he or she is associated with, no matter how weak this association is, (2) *impersonal causation* or commission, whereby the actor is perceived to be responsible for outcomes he or she caused by committing an act, even if these outcomes were not foreseen and therefore not intended, (3) *foreseeability*, whereby the actor is perceived to be responsible for outcomes which he or she could anticipate, even if they were not intended, (4) *personal causation*, whereby the actor is perceived to be responsible for outcomes he or she intended, and (5) *justifiability*, whereby the actor is perceived to be responsible for outcomes which cannot be justified by the situation the actor found him- or herself in (such a situation being acting in self-defence, for example). Accountability would seem to increase from the first to the fourth level, with the fifth level pointing to mitigating circumstances. Hamilton (1978) presented a model which re-interpreted each of Heider's levels in terms of legal responsibility rules. The first level, for example, is equivalent to vicarious liability (e.g., when a bar-owner is held responsible for the sale of alcohol to individuals under age on his or her premises, regardless of whether he or she is aware of it and has agreed to it, Semin & Manstead, 1983), whereas the fourth level represents full responsibility in the sense of the law.

Weiner (1995a, b) suggests that the process of attributing responsibility follows a certain sequence. First, an observer determines whether there is human involvement in the event, that is, whether there is *personal* as opposed to impersonal *causality*. Weiner gives the example of a car being damaged either by another car hitting it or by a falling rock. Only if the driver of another car was involved, would the observer continue with the responsibility inference process; according to Weiner (1995a), "assignment of responsibility requires human or personal agency" (p. 6). The next step involves the observer deciding whether the actor

was able to *control* the event. If the driver of the other car was not able to control his or her behaviour, for example, because of having fainted, he or she will not be regarded as responsible. Finally, the observer checks for any *mitigating circumstances*, such as other morally valued goals which justify the behaviour, or an inability to comprehend the 'wrongness' of an act on the part of the actor. If mitigating circumstances are found to have been present, responsibility can be diminished or entirely lifted.

For a judgement or expectation of personal responsibility as an antecedent to accountability it seems unlikely that mitigating circumstances play an important role, because mitigating circumstances are typically brought forward *after* a person has been called to account for his or her behaviour. Weiner (1995a, b) recognises that his proposed sequence is not invariant, for example, people may arrive at premature conclusions which are revised as more information becomes available. Mitigating circumstances appear to be particularly likely to be left out of the inference process, since it is difficult to assess them without a detailed knowledge of the situation. Also, an actor may recognise that others may not be aware of mitigating circumstances and therefore he or she will expect to be held responsible even if mitigating circumstances are present. Whereas Heider (1958) explicitly assumes that there is no difference in the process of responsibility attribution for an external observer and oneself as long as the same information is available, it remains unclear whether Weiner proposes attribution of responsibility to oneself to follow the same sequence as the one suggested for an external observer. For the purposes of this discussion, however, anticipation on part of the individual that others would consider him- or herself personally responsible for his or her actions seems sufficient to bring about the intended effects.

When comparing Heider's concept of personal causation with Weiner's concept of personal causality, it becomes apparent that, despite their similar name, these two concepts are quite different. Heider's personal causation refers to bringing about intended outcomes, whereas Weiner's personal causality only implies personal involvement, but not necessarily control over and/or an intention to bring about a certain outcome. Weiner's concept of personal causality, therefore, resembles Heider's concept of *im*personal causation. Weiner (1995a) does discuss intention as a factor that influences perceptions of responsibility, but sees it as a determinant of the *degree* of responsibility (if an act was intended, responsibility is higher) rather than a necessary condition of perceptions of responsibility.

We will go along with Weiner's suggestion and assume that two factors are crucial for a perception of personal responsibility as an antecedent of accountability: *personal causality* (as understood by Weiner) and *control* over an event. Note that these two factors also are part of Schlenker's (Schlenker, Weigold, & Doherty, 1991; Schlenker, Britt, Pennington, Murphy, & Doherty, 1994) triangular model of accountability, as two variables determining the identity-event linkage.

Evaluation apprehension

Clearly, anticipated evaluation is a central feature of accountability. The individual expects to be evaluated by other people and/or, in some cases, carries out a self-evaluation. Many definitions of accountability, indeed, only focus on the evaluation aspect and neglect the aspect of explanation or justification. Examples are the definitions by Carnevale (1985), who describes accountability as the concern about the evaluation of one's views and preferences by others, or Yarnold, Mueser and Lyons's (1988) view of accountability as a public personal evaluation.

The effects of evaluation apprehension have been empirically investigated in many different domains, but have received particular attention in studies on the effects of the presence of other people on task performance. These studies have some relevance here, as the performance of accountable decision makers is under investigation.

Evaluation apprehension is a result of the presence of others and can have both positive and negative effects on task performance. It enhances performance if the task is easy (social facilitation), and debilitates performance if the task is difficult (social impairment, 'choking under pressure', Baumeister, 1984). Zajonc (1965) proposed that the mere presence of others causes an alertness response which is accompanied by a heightened physiological arousal. This arousal increases the likelihood of dominant responses (simple, well-learned behaviours that are highest in the habit hierarchy) and decreases the likelihood of non-dominant responses (novel or complex behaviours). If the dominant response is the correct or appropriate one, as is the case for an easy task, the presence of others will enhance performance. If the dominant response is incorrect or inappropriate, for example, when the task is difficult, the presence of others will impair performance. Cottrell (1972), however,

suggested that the drive-like effects of the presence of others come about because the individual has learned that others are a source of positive or negative evaluation and that the presence of others often co-varies with rewards or punishments. The anticipation that other people will *evaluate* the individual's performance, that is, evaluation apprehension (and not their mere presence, as Zajonc suggested), according to Cottrell, increases the individual's level of arousal and causes the effects described above. A study by Markus (1978) designed to test between these two approaches suggested that while mere presence seems to be sufficient to elicit the typical performance effects, evaluation apprehension has an effect over and above the effect of the mere presence of other people.

Self-presentational accounts, such as impression management theory (Tedeschi & Riess, 1981), assume that evaluation apprehension is a fear of not being able to present a competent public image. Making a mistake when other people are present leads to embarrassment. This embarrassment is assumed to be an impediment to cognitive and motor control and impairs performance, especially the performance of those behaviours where a failure is likely or has occurred before (Bond, 1982). In this account, task complexity only plays a role in so far as it conveys information about the performer's task competence. Success in easy tasks signals to the individual that the desired public image can be maintained, whereas failure in difficult tasks signals that the desired image cannot be maintained and causes embarrassment. This suggests that performance in the presence of other people should not so much depend on the actual difficulty of the task as on the individual's perception of his or her own competence at the task, mediated by the perceived difficulty of the task. In line with this hypothesis, Bond (1982) was able to demonstrate that the presence of an observer impaired the learning of simple items when those items were embedded within a difficult task and did not impair the learning of complex items when they were embedded within an easy task.

These last results underline the importance of self-presentational and identity-related considerations for accountability effects, as suggested by Schlenker and colleagues (Schlenker & Weigold, 1989; Schlenker, Weigold & Doherty, 1991; Schlenker, Britt, Pennington, Murphy, & Doherty, 1994) and Semin and Manstead (1983). According to Schlenker et al., accountability is assumed to pose a threat to identity and the strategies applied to cope with this threat have repercussions on the identity of an individual. Identity is

understood as a theory of the self about "how one is and should be perceived, regarded and treated in social life" (Schlenker, 1982, p. 194). It is both a structure that contains the ideas about what one is like and an active process presenting oneself to be a particular kind of person (Schlenker, 1986). This process, Schlenker refers to as 'self-identification'. The process of self-identification links identity with accountability, as the construction and evaluation of identity takes place within the context of accountability. Similarly, Semin and Manstead (1983) argue that the attribution of motives to an individual's actions means that the actor is regarded as possessing an identity which is defined by those actions. When an individual's conduct is called into question, the issue of identity management therefore becomes salient.

The implied presence of others and identifiability as necessary antecedents of evaluation apprehension.

One condition that is clearly necessary for the individual to feel evaluation apprehension is the implied presence of others. Some authors, for example, Semin and Manstead (1983) and Tetlock (1992) have argued that it may be sufficient for this presence to be imaginary rather than real, as individuals have internalised the external behavioural rules to which they want to adhere and then engage in a process of self-evaluation. An evaluation by others can only be carried out, however, if the behaviour that is to be evaluated is identifiable (Harkins, 1987; Harkins & Szymanski, 1987). It therefore seems reasonable to assume that in order for individuals to feel evaluation apprehension, they must also know that their behaviour and its potential consequences are identifiable to others.

Identifiability has been considered an important explanatory variable in social loafing and de-individuation research. Social loafing is the phenomenon of individuals reducing their effort when they work in groups compared to when they work alone, because they think that the other group members will expend the necessary effort (Latané, Harkins & Williams, 1979; see also Kerr and Bruun, 1983, about the 'free-rider effect'). Similarly, de-individuation is a state of lack of self-awareness in which the individual, as a consequence of being submerged

in a group, shows more disinhibited, often anti-normative behaviour[1] (Festinger, Pepitone, & Newcomb, 1952; Zimbardo, 1969, 1970).

Both social loafing and de-individuation occur because individual behaviour is not identifiable. In the case of social loafing, the anonymity of pooled performance diffuses responsibility for performance and protects individuals from any negative effects of not exerting the same effort as the other group members. Individuals who loaf can hide in the crowd, but at the same time they can also become lost in the crowd; they cannot receive credit for their positive contributions. If people's individual output can be identified, social loafing disappears (Jackson & Harkins, 1985; Williams, Harkins, & Latané, 1981). Similarly, the lack of identifiability in de-individuation situations (in experiments often manipulated by giving participants the opportunity to hide their identity behind masks, e.g., Zimbardo, 1969) allows individuals to escape from any negative effects of anti-social behaviours.[2]

Salient norms as a consequence of evaluation apprehension and personal responsibility

In the previous chapter, accountability was described as driving both social control and self-regulation by enforcing normative behaviour. In the following, we will describe the way in which the two main antecedents of accountability, evaluation apprehension and personal responsibility, are linked to this norm-enforcement mechanism.

Cialdini, Kallgren and Reno (1991) distinguish between descriptive norms and injunctive norms. Descriptive norms refer to what is commonly done or what is normal, whereas injunctive norms refer to what is either commonly approved of or what is socially sanctioned. These two types of norms usually coincide, but not always. For example, a typical phenomenon observed in emergency situations is the bystander effect, whereby a person in

1 Reicher (1987), however, has recently argued that de-individuation does not reduce the influence of normative constraints on the individual's behaviour, but increases the feelings of shared group membership and the power of group norms, thereby increasing whatever behaviour is typical for the group.

2 Prentice-Dunn and Rogers (1982) have contested this assumption. According to their differential self-awareness theory, it is an external focus of attention rather than a lack of identifiability that results in de-individuation.

need of help is increasingly less likely to receive help the more bystanders are present (e.g., Latané & Darley, 1970). The injunctive norm in this situation would be to step in and help. However, because the other bystanders are passive, individuals infer a descriptive norm which suggests that intervention would be inappropriate, and do not act, because they do not want to embarrass themselves in front of other people. The effect may be strengthened by the existence of another, injunctive norm to mind one's own business. If one person offers help, on the other hand, he or she will serve as a model, defining help as an appropriate response, and others will be quick to follow (Staub, 1978).[3] The significance of descriptive norms is also reflected in Deutsch and Gerard's (1955) concept of informational social influence, which is social influence that occurs because the individual is unsure about what constitutes appropriate behaviour and looks to other people to see how they behave.

Both descriptive and injunctive norms provide external standards of appropriate behaviour. As mentioned above, however, external norms often become internalised and turn into personal norms. Schwartz (1977) defines personal norms as self-based standards or expectations for behaviour which are enforced through processes of self-enhancement or self-deprecation; hence, not only the standards but also the sanctions for not obeying them are located within the self. While many studies have suggested that personal norms are better predictors of behaviour than external social norms (see Schwartz, 1977, for a review), there has, however, also been evidence that if a social norm is activated, its influence on behaviour can override that of personal norms. An example is a study by Cialdini, Kallgren, and Reno (1991, study 9), which showed that focusing participants' attention on an injunctive social anti-littering norm resulted in significant less subsequent littering compared to participants in a self-focus condition, irrespective of whether participants' personal anti-littering norm was weak or strong. Also, littering was significantly lower than in a condition with a social evaluation focus that was unrelated to the relevant normative behaviour, suggesting that the possibility of social evaluation per se was not sufficient to affect participants' behaviour.

[3] Interestingly, descriptive norms have been found to be asymmetrically effective in that they are more likely to influence behaviour in a negative than in a positive way (Cialdini, Kallgren, & Reno, 1991).

These results, apart from illustrating the potentially powerful influence of social norms on behaviour, underline the importance of norm activation for any such influence. Norms only influence behaviour if they are activated, that is, if they are made salient or are focused upon.

Self-attention theories, for example, that by Carver and Scheier (1981a, b; see also Duval & Wicklund, 1972) have suggested that the presence of an evaluative audience makes behavioural norms salient and instigates a process of 'matching to standard', that is, enhanced conformity to such norms (see also Heider, 1958). In their cybernetic theory of self-regulation, Carver and Scheier propose that the presence of other people (as well as other conditions, such as the presence of a mirror or a camera) heightens an individual's self-directed attention. This, in turn, causes the individual to engage in a behavioural feedback loop, similar to the Test-Operate-Test-Exit (TOTE) unit suggested by Miller, Galanter, and Pribram (1960), whereby the individual's present behaviour is compared to a salient standard of behaviour. If a discrepancy is detected, measures are taken to reduce the discrepancy. If a subsequent comparison reveals that the discrepancy still exists, the process is repeated until no discrepancy can be detected anymore[4]. A heightened self-focus thus leads to enhanced conformity to a salient standard. The standard is usually inferred from situational cues. Depending on the nature of these cues and how the individual interprets them, different standards may become salient and different behaviours may result. Grush (1978) could show that participants playing a Prisoner's Dilemma Game in front of an audience described as previous winners at the game made more competitive choices than no-audience control participants, whereas participants playing the game in front of an audience described as 'Human Relations Council' made fewer competitive choices than control participants. This was the case despite all participants having been given a pre-treatment experience with the game causing them to favour competitive choices.

Whereas in Carver and Scheier's (1981a, b) model of audience effects self-directed attention is assumed to be the central mechanism that mediates audience effects, other models like that of Zajonc (1965) or Cottrell (1972) see physical arousal as the central mediator. In

[4] The 'matching to standard' process which reduces the discrepancy between actual and normative behaviour is only instigated, however, if the individual's expectancy of being able to reduce the discrepancy is sufficiently favourable. If this is not the case, the individual is likely to withdraw, either physically or mentally.

the latter case, one would still expect relevant norms to become salient. Easterbrook (1959), for example, demonstrated that physical arousal causes a perceptual narrowing, resulting in a focus on those cues that appear most relevant.

It may be argued that not only evaluation apprehension but also personal responsibility is likely to increase the individual's self-directed attention and thereby make certain norms salient, as it draws attention to the possible consequences of one's behaviour. However, unlike evaluation apprehension, personal responsibility is likely to make personal rather than external norms salient, as there is no anticipation of being evaluated by others and behaviour should therefore flow from one's own values. It is therefore assumed that personal responsibility will make the individual focus on his or her personal norms and encourage a matching to personal standards, whereas evaluation apprehension will make the individual focus on possible audience norms and encourage matching to perceived audience standards.

Summary and conclusions

The previous discussion attempted to shed some light on the antecedent conditions of accountability. Two processes were identified as being central to generating accountability pressures in the individual, a feeling of personal responsibility for performing a behaviour that has potentially negative consequences, and the expectation of being evaluated. Drawing on attribution research, in particular on models put forward by Heider (1958) and Weiner (1995a, b), it was concluded that in order for the individual to feel personally responsible, two conditions have to be met, (1) personal causality as defined by Weiner (1995a, b), that is, the individual must have brought about an outcome by his or her behaviour, although the outcome may not have been intended, and (2) control over the behaviour. The second accountability component, evaluation apprehension, was also seen to be based on the fulfilment of two conditions, (1) the implied presence of others, and (2) identifiability of the actor's behaviour. It was argued that both personal responsibility and evaluation apprehension make certain behavioural norms or standards salient that the individual will strive to adhere to. In the case of personal responsibility, internal or personal norms will be made salient, whereas in the case of evaluation apprehension, external or audience norms will become salient.

The results of studies addressing evaluation apprehension, although not directly bearing on the effects of accountability on behaviour, allow some interesting predictions. Firstly, evaluation apprehension research suggests that accountability should be accompanied by a heightened physiological arousal that mediates its effects. To our knowledge, no accountability study exists where physiological measures have been taken to show that this is indeed the case. Secondly, task difficulty should emerge as one of the variables that moderate accountability effects. This, again, has not yet been explicitly investigated but one would expect that if the task is easy, accountability should enhance performance, whereas if the task is difficult, accountability should result in a decreased performance (see also Pelham & Neter, 1995). For testing the validity of this hypothesis, the particular operational definition of task difficulty that is employed seems to be crucial. Evaluation apprehension studies have related task difficulty to the appropriateness of dominant and non-dominant responses, but the subjective perception of a task as being difficult has also been shown to play an important role. The evidence for accountability enhancing dominant responses is mixed at best (see Lerner & Tetlock, 1999) and whether or not easy tasks in this case can be equated with those for which dominant (i.e., habitually strong) responses are likely to be correct is questionable. It may be argued that accountability is likely to draw attention to the expectations of the audience the individual is accountable to, and that task difficulty is at least partly dependent on whether the perceived norm against which performance is to be measured can guide the individual's behaviour and provide cues as to how a positive evaluation can be achieved. If this is not the case, the task is likely to be perceived as difficult and the expectation to be evaluated is likely to induce stress and decrease performance (see Siegel-Jacobs & Yates, 1996). This issue will be discussed in more detail in the next chapter. It is important to keep in mind, though, that accountability cannot be equated with evaluation apprehension, as the latter is just one of the central components that generate accountability pressures.

Along the same lines, it must be stressed that, despite their superficial similarity, an accountability manipulation is not identical with placing participants in a transmission set, that is, leading them to expect that they will have to communicate their opinions or impressions of an event to others (as opposed to a reception set, where they receive additional information from others) (Tetlock, 1992). Whereas placing participants in a transmission set focuses their attention on producing an easy-to-understand, logically consistent description,

making participants accountable additionally implies that they will not only have to communicate their opinion but also defend it against possible counter-arguments.

An important question that has not been addressed yet relates to the necessity and/or sufficiency of the conditions specified above. In order to answer this question, we must first refer to the distinction between external and internal accountability. *External accountability* is a demand to justify one's behaviour to others, whereas *internal accountability* is a need to justify one's behaviour to oneself (Simonson, 1989). External accountability can either be implicit, when a person feels observed by other people and has a desire to appear competent, without actually having been challenged by others to explain his or her behaviour, or it can be explicit, when individuals are told that they are accountable to someone else (Simonson, 1989). Explicit external accountability is present when formal accountability procedures are in place, for example, in organisations where employees have to report to their managers and are aware of this. Explicit external accountability may also be present in negotiation contexts when the leaders of the negotiating parties as agents make a decision about a settlement on behalf of other people, their principals (Pratt & Zeckhauser, 1985). The main difference between external and internal accountability is the source of behavioural control. In the case of external accountability, control is exerted by the person(s) the individual is accountable to, whereas internal accountability implies that control lies within the individual him- or herself. Internal accountability arises because decision makers tends to internalise the criteria employed by others and want to prove to themselves that they have behaved in an appropriate manner.

It is argued here that different combinations of the above-mentioned antecedent conditions result in either internal or external accountability (see Figure 2.1). In particular, a feeling of internal accountability is hypothesised to occur when the individual has personally caused or is about to cause a potentially negative outcome by behaviour that is/was under his or her control, resulting in a feeling of personal responsibility for this outcome. Personal causation and control are both considered necessary antecedents of personal responsibility which, in turn, is both necessary and sufficient to create internal accountability. External accountability, on the other hand, will only be induced if the individual, in addition to feeling personally responsible, expects to be evaluated by others. The implied presence of others and identifiability of the individual's behavioural outcomes to them are necessary conditions for

evaluation apprehension (see also Weldon & Gargano, 1988). Evaluation apprehension and personal responsibility are both necessary but not sufficient conditions for external accountability.

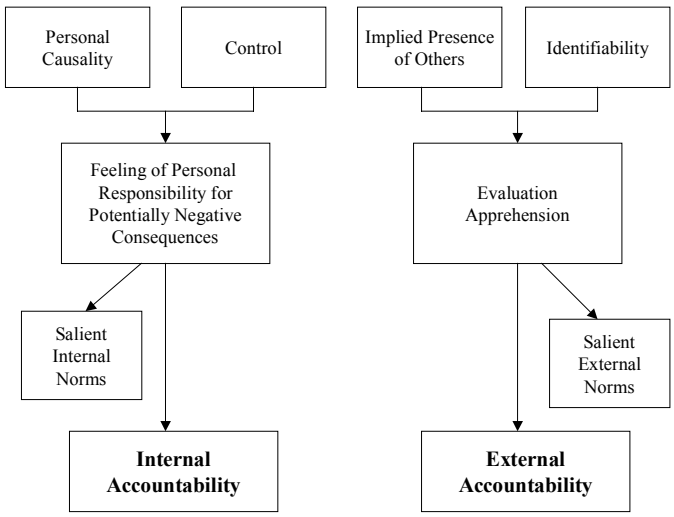


Figure 2.1. The antecedent conditions of internal and external accountability.

Having developed an initial model of the antecedent conditions of both internal and external accountability, the following chapter will provide a review of studies that have investigated the effects of internal and external accountability, with the ultimate goal of further developing the model such that certain antecedent conditions can be linked to certain accountability effects.

CHAPTER 3: THE EFFECTS OF INTERNAL AND EXTERNAL ACCOUNTABILITY ON DECISION PROCESSES AND OUTCOMES

This chapter will present a review of the existing literature on the effects of internal and external accountability on decision processes and outcomes. It will be argued that, generally, accountability has two main effects on decision processes, namely, (1) it increases the amount of information that is considered and the complexity with which this information is processed, and (2) it increases the decision maker's motivation to arrive at a decision that is easy to justify. In order to explain the differential effects of accountability, three classifications of different types of accountability will be presented, (1) accountability to an audience with known vs. unknown views, (2) pre-decisional vs. post-decisional accountability, and (3) procedural vs. outcome accountability. Finally, the effects of accountability will be linked to its antecedent conditions of personal responsibility and evaluation apprehension, as described in the previous chapter.

The effects of internal accountability

Research on the effects of accountability on behaviour in general and decisions in particular has rarely focused on internal accountability. It may be argued, however, that studies intended to test the theory of cognitive dissonance by Festinger (1957) often have implicitly tapped into internal accountability processes (Staw, 1980). The antecedent conditions of cognitive dissonance are very similar to those of internal accountability, namely a feeling of personal responsibility for foreseeable, potentially aversive consequences in a situation of free choice (see Cooper & Fazio, 1984, for a review), resulting in a need for self-justification. Studies investigating the effects of cognitive dissonance in the context of decision-making have typically found that after a decision, the chosen alternative is evaluated better than before, thereby reducing the dissonance that arises when one has to put up with the disadvantages of the chosen alternative and loses the advantages of the non-chosen alternatives (e.g., Brehm, 1956). Also, a selective exposure effect may be observed, where people prefer to expose themselves to information that supports their chosen alternative and

avoid information that contradicts it (e.g., Frey, 1986). Although not anticipated by Festinger himself, some studies have shown that biased information processing can already be observed after a tentative decision but before a final decision has been made (e.g., Borcherding, 1981; Grabitz & Haisch, 1972). [1]

Staw (1976) investigated the effects of personal responsibility on participants' tendency to escalate their commitment to a failing course of action. Perceived personal responsibility, it has been argued in the previous chapter, is a necessary antecedent condition of internal accountability. In this case, however, personal responsibility was not prospective but retrospective. Participants were asked to allocate R&D funding to one of two corporate divisions of a fictitious company. In the 'high personal responsibility' condition, participants were asked to make this decision twice and received either positive or negative feedback on the outcome of their initial decision, in the 'low personal responsibility' condition the initial (either successful or unsuccessful) decision was said to have been made by someone else. It was found that if participants perceived high personal responsibility for the previous decision and this decision had resulted in negative consequences, they invested a disproportionately high amount of money in their previously chosen alternative, suggesting that they felt a need to justify their choice to themselves and defend themselves against the negative consequences resulting from it.

Weldon and Gargano (1985) tested the effect of prospective personal responsibility on cognitive effort. Responsibility for the judgement task, an evaluation of a series of part-time jobs described on five dimensions, allegedly carried out for a (fictitious) vocational research institution, was manipulated by either telling participants that they were the only evaluator of their particular set of job descriptions ('individual responsibility' condition), or that they were 1 of 16 students evaluating that particular set of job descriptions ('shared responsibility' condition). These instructions clearly established personal responsibility in the former and an absence of personal responsibility in the latter condition, thereby implicitly manipulating internal accountability. It was found that participants under individual responsibility produced

1 A more detailed discussion of dissonance theory and the effects of cognitive dissonance on information search and evaluation will follow in Chapter 4.

more evaluations and used more complex judgement strategies that were better predicted by a linear additive model than participants in the shared responsibility condition.

Another line of research that may be said to have investigated internal accountability, albeit not explicitly, is research on information processing under different degrees of personal involvement. In Petty, Cacioppo, and Schumann's (1983) study, for example, participants were asked to report their attitude towards a new brand of disposable razor after having been exposed to an advertisement for this product. High involvement was created by telling participants that the product would soon be available in their home town and that, as a reward for taking part in the study, they would later be allowed to choose among several brands of disposable razors. The latter information, it may be argued, made personal consequences of their performance salient to the participants and thereby created internal accountability. In the 'low involvement' condition, participants were told that the product would soon be tested in some other towns and that they would be allowed to choose among several brands of toothpaste as a reward for taking part in the study. The advertisement contained either strong or weak arguments and the product was either endorsed by a famous sports celebrity or by an unknown ordinary citizen. It was found that 'high involvement' participants processed the message in a more systematic way than 'low involvement' participants, indicated by the fact that their attitude toward the advertisement only depended on the strength of the arguments presented and not on the famousness of the endorser. Similarly, Borgida and Howard-Pitney (1983) found that personal involvement, that is, being personally affected by an issue, eliminated the influence of a discussant's visual salience on participants' ratings of her influence and causal role and resulted in a more systematic processing of message arguments.

These findings confirm Cacioppo and Petty's (1981, 1986) elaboration likelihood model which suggests that highly involved participants take a central processing route, where they carefully think about arguments and therefore are influenced by the quality of the presented arguments, whereas low-involvement participants take a peripheral processing route, where they rely on heuristics to evaluate a message superficially and therefore are influenced by peripheral cues such as communicator likeability, salience, or expertise.

In Harkness, DeBono, and Borgida’s (1985) study, which tested the effects of personal involvement in a different context, participants were asked to perform a co-variation

judgement task in which they were presented with 2 x 2 co-variation matrices containing information about a person's dating choices. High involvement was created by telling participants that they would date the person later (thereby making personal consequences and responsibility for the consequences of their performance salient), in the 'low involvement' conditions, participants either believed that they would date a different person or that the presented information was fictitious. The co-variation judgements of 'high involvement' participants were consistent with the use of more complex strategies and tended to be more accurate than those of 'low involvement' participants.

Lee, Herr, Kardes, and Kim (1999, exp. 2) directly compared the effects of issue involvement and external accountability on information search and use in a multiattribute decision task. Participants were asked to choose one of nine laptop computers, each described by eight attributes on a computerised information board. For half of the participants, high issue involvement was induced by telling them that their Business College was planning to purchase laptop PCs to lend them to students and that their decision would help to select a suitable model for them. The other participants in the 'low involvement' condition were told about the same kind of scenario, but it was said to apply to another Business School, not their own. External accountability was then manipulated by reminding half of the participants about the importance of a correct choice and telling them that they would be required to write a justification of their choice afterwards (high accountability), whereas the other half was told that the study was not about assessing the correctness of their choice but identifying individual preferences (low accountability). Results showed that *either* high issue involvement *or* high accountability were sufficient to make participants examine more information, spend more time examining information and engage in more compensatory information processing than participants in the other conditions. Where the effects of the two manipulations differed was in terms of their contingency on prior knowledge levels. Whereas high accountability always increased the depth and complexity of information processing, no matter whether participants' previous knowledge about laptop computers was high or low, high issue involvement only increased search effort and compensatory processing if participants had little prior knowledge. Lee et al. (1999) explained this effect by assuming that high external accountability focuses participants' attention on extrinsic rewards, such as other people's approval of their decision, which is unlikely to be gained unless much effort is

invested in the task, whereas high issue involvement focuses participants' attention on intrinsic rewards, such as a feeling that they have made a good decision. The latter can be gained with relatively little effort unless participants lack sufficient experience, which they will then attempt to compensate for by increasing effort. Note that this explanation is consistent with the point made at the end of the previous chapter, namely, that personal responsibility as an antecedent of internal accountability will make personal norms salient, whereas evaluation apprehension as an antecedent of external accountability will make external norms salient.

A comparison of the effects of personal responsibility and external accountability was also the focus of Rozelle and Baxter's (1981) study. Participants had to rate mock applicants to graduate school on a series of self-and experimenter-generated traits. Personal responsibility was manipulated by either telling participants that their ratings would be considered along with other materials when reaching a decision about whether to accept or reject the applicants ('high personal responsibility' condition), or telling them that the selection decisions had already been made and their ratings were combined with other materials in order to improve the selection process in coming years ('low personal responsibility' condition). Participants were also either made externally accountable by being led to believe that they would have to discuss their ratings with other judges and that the applicants would be able to inspect their ratings ('high accountability' condition), or they were not made accountable, in which case no mention was made of any group discussion and participants were told that their ratings would be kept in strict confidence and applicants would not be able to inspect them. As in Lee et al.'s (1999) study, results suggested that either high personal responsibility or high external accountability were sufficient to increase between-judge agreement on the same target, decrease between judge-agreement on different targets and decrease within-judge agreement on different targets, suggesting that participants' ratings under these conditions more reliably reflected real characteristics of the target rather than perceptual idiosyncrasies of individual judges. The effect of personal responsibility on these measures was not as strong as that of external accountability, however.

A final study comparing the effects of personal responsibility and external accountability to be reported here is that by Weldon and Gargano (1988). The cover story and manipulation of personal responsibility was identical to that of their 1985 study described

above. Additionally, half of the participants were told that the experimenters may want to contact them afterwards, in order for participants to explain what information they used and why they made certain judgements ('accountability' condition). In the 'no accountability' conditions, no mention was made of these things. Note that the manipulation of external accountability contains an implicit suggestion of identifiability. As Weldon and Gargano point out, this was intended, as they, in line with the analysis presented in the previous chapter, assume that identifiability, despite being distinct from external accountability, is a necessary component of it. The results showed a significant interaction effect between personal responsibility and accountability on cognitive effort, such that the number of non-zero beta weights in a linear additive regression model, as indicators of how much information had been used, was greater for participants who believed they were the only evaluator than for participants who believed they were 1 of 16 evaluators, whereas there was no difference when participants had been made accountable. Also, for participants who thought they shared responsibility for the judgement task (regardless of whether they had been made externally accountable or not), a smaller proportion of judgement variability could be explained by a linear regression model, suggesting that their judgement strategies were less linear. This latter result may also be interpreted in terms of consistency, as an inconsistent application of a linear combination strategy (or random evaluations, in an extreme case) would also result in a low proportion of explained variance. Given that this result was obtained even when shared responsibility participants were accountable, Weldon and Gargano (1988) concluded that accountable participants only expected their performance to be judged with regard to the information they had used (as this was explicitly mentioned in the instructions), whereas they did not expect their consistency to be judged. This again supports the idea that particular norms become salient, when the individual experiences evaluation apprehension.

In conclusion, the findings of the studies reviewed above suggest that internal accountability, operationalised in the form of personal responsibility or high personal involvement, increases the depth of information search and complexity of information integration. Under certain circumstances, however, in particular, after a decision has been made, it seems to increase people's tendency to bolster their previous choice by searching for information that is consistent with it, by changing their evaluation of their chosen alternative

in such a way that its attractiveness is increased, and by continuing to invest in it even though it has been proven wrong.

The effects of external accountability

Research on how external accountability affects decision processes generally has also produced these two kinds of results. On the one hand, it has been found that external accountability, compared with no accountability, results in more thorough and multi-dimensional, that is, more analytical and compensatory information processing. On the other hand, it has been found that external accountability leads to a greater defensiveness and to decisions which are more prone to a tendency to choose according to the ease of justifiability instead of making the best decision.

Increased effort and complexity of information processing

A higher vigilance and complexity of information processing when the individual is externally accountable compared to when he or she is not accountable, have been indicated by

- a deeper information search (Billings & Scherer, 1988; Doney & Armstrong, 1996; Hattrup & Ford, 1995; Huber & Seiser, 2001)
- a longer time to arrive at a decision or spent on a task (Doney & Armstrong, 1996; Ford & Weldon, 1981; Hattrup & Ford, 1995; Huber & Seiser, 2001; McAllister, Beach & Mitchell, 1979, exp. 3; Mc Lean Parks & Conlon, 1990; Siegel-Jacobs & Yates, 1996; Stewart, Billings, & Stasser, 1998; Van Hiel & Schittekatte, 1997; Weldon & Gargano, 1988)
- an increased self-reported task motivation (Doney & Armstrong, 1996; Magjuka, Baldwin, & Lower, 1994)
- a higher integrative complexity of thoughts (Gordon & Stuecher, 1993; Koonce, Anderson, & Marchant, 1995; Mandel, Axelrod, & Lehman, 1993; Tetlock, 1983a; Tetlock & Boettger, 1989; Tetlock & Kim, 1987; Tetlock, Skitka, & Boettger, 1989)

- the application of more complex, compensatory decision rules (McAllister, Beach & Mitchell, 1979, exp. 3; Murphy, 1994)

Huber and Seiser (2001), for example, asked their participants to role-play a member of a social services executive board with the task to choose a manager for a home for criminal adolescents. Participants were either only told that they had been delegated by the executive board to choose the candidate ('no justification' condition), or were additionally told that, after the decision, they would have to write a letter to the executive board in which they would either have to explain why the selected candidate was chosen ('accounting' condition) or they would have to convince the other members of the executive board to follow their recommendation ('convincing' condition). Six candidates, described on 16 attributes, were presented on an information board. Apart from standard information board measures, such as decision time and percentage of utilised information, think-aloud protocols were analysed. The results showed that decision time was longest in the 'convincing' condition and shortest in the 'no justification' condition, with the 'accounting' condition in-between. Also, the percentage of utilised information was larger in the justification conditions than in the 'no justification' conditions. No significant difference was found between the 'accounting' and the 'convincing' condition, however. The think-aloud protocols, on the other hand, did show a significantly larger number of statements in the 'convincing' than in the 'accounting' condition. Again, the lowest number of statements was observed in the 'no justification' condition. This study therefore demonstrates increased task effort as a consequence of external accountability pressures. Increased task effort has also been reported by participants directly, for example, in a study by Magjuka, Baldwin, and Loher (1994), in which participants in a business skills training programme were either made externally accountable or not made accountable for their performance.

Tetlock and his co-workers (Tetlock, 1983a; Tetlock & Boettger, 1989; Tetlock & Kim, 1987; Tetlock, Skitka, & Boettger, 1989) have repeatedly demonstrated a higher integrative complexity of thoughts, when individuals are made externally accountable (see also Mandel, Axelrod, & Lehman, 1993), particularly when they are made accountable to an audience with unknown views. Tetlock (1983a), for example, asked participants to report their thoughts and respond to a series of attitude scales on three controversial social issues, affirmative action, capital punishment, and increased defence spending. This happened under

one of four conditions, (1) no accountability, (2) accountability to someone with liberal views, (3) accountability to someone with conservative views, or (4) accountability to someone with unknown views. Participants' thoughts were coded for their integrative complexity, that is, the extent to which participants focused on multiple dimensions of a problem and the development of complex connections among these multiple dimensions. Integrative complexity was found to be higher when participants had been made accountable to an individual with unknown views than when they had not been made accountable. Interestingly, the lowest integrative complexity was obtained when participants had been made accountable to an individual with either liberal or conservative views. This suggests that when participants know or can anticipate the view of the audience they have been made accountable to, they will just adopt this view and only expend minimal cognitive effort.

Studies which have analysed participants' decision strategies rather than their thoughts on a topic have found that external accountability results in the employment of more complex decision strategies. Murphy (1994), for example, showed that when participants were told that they would have to justify their responses to the experimenter and other participants, they used more complex strategies when assessing the co-variation of two cues. McAllister, Mitchell, and Beach (1979, exp. 3) asked participants, who had either been made externally accountable or had not been made accountable for the particular decision task they were faced with, to make a decision on the potential marketability of sports products[2]. In order to arrive at a decision, participants had to choose one of four different prescribed decision strategies which differed in the amount of information they used, the difficulty of computations involved and their past record for accuracy. Participants who had been made accountable were found to select more analytical decision strategies than participants who had not been made accountable. In addition, they allocated more time to and actually spent more time on completing the task. De Hoog and van der Wittenboer (1986), in a similar study, were not able to replicate these results, however. Participants who had been told to justify their decision rule to a panel of experts were not more likely to select compensatory decision rules (the averaging rule and least objections rule) than participants who had not been made

[2] In addition to accountability, perceived task significance and decision reversibility were manipulated. Since no interactions of these factors with the accountability factor were observed, the effects of these manipulations are not reported.

accountable. De Hoog and van der Wittenboer suggest that a possible explanation for their failure to find a difference between accountable and non-accountable participants' rule selection may have been participants' low familiarity with the particular non-compensatory decision rules offered. Generally, the procedure of having participants choose pre-defined decision rules is very artificial and may not reflect participants' natural use of decision rules anyway.

The increased cognitive effort that accountability seems to instigate has a number of beneficial effects, for example,

- a greater consistency in the application of judgmental strategies (Ashton, 1992; Hagafors & Brehmer, 1983; Ordonez, Benson III, & Beach, 1999)
- a higher self-awareness of decision makers in terms of how they use information to arrive at their decision (Cvetkovich, 1978; Johnson & Kaplan, 1991)
- a higher consensus among decision makers (Ashton, 1992; Rozelle & Baxter, 1981; Johnson & Kaplan, 1991)
- a greater accuracy of judgements and decisions (Ashton, 1992; Cvetkovich, 1978; Nero & Motowidlo, 1995; Rozelle & Baxter, 1981) or, more generally, a better performance (Magjuka, Baldwin, & Lower, 1994; Tausky & Chelte, 1983; Yarnold, Mueser, & Lyons, 1988)
- a better recall of information (Chaiken, 1980, exp. 1; Tetlock, 1985b)
- a de-biasing effect, particularly for biases which result from incomplete information processing (Bodenhausen, Kramer, & Süsser, 1994; Boudreau, Baron, & Oliver, 1992; Brockner, Shaw, & Rubin, 1979; Foschi, 1996; Freund, Kruglanski, & Shpitzajzen, 1985, exp. 3; Kassin, Rigby, & Castillo, 1991; Kennedy, 1993; Kirby & Davis, 1998; Kroon, 't Hart, & van Kreveld, 1991; Kroon, van Kreveld, & Rabbie, 1992; Kruglanski & Freund, 1983; Lerner, Goldsmith, & Tetlock, 1998; McKenna & Myers, 1997; McLean Parks & Conlon, 1990; Mosier, Skitka, Heers, & Burdick, 1998; Schadewald & Limberg, 1992; Simonson & Nye, 1992; Slugoski, unpublished; Takemura, 1993; Tetlock, 1983b,

Tetlock, 1985b; Tetlock & Kim, 1987; Webster, Richter, & Kruglanski, 1996; Wells, Petty, Harkins, Kagehiro, & Harvey, 1977)

Studies which have employed a multiple cue prediction (or policy capturing) paradigm have repeatedly shown that external accountability improves various aspects of the judgement process. Such studies typically present participants with at least two information cues and ask them to make a prediction based on these cues. In Hagafors and Brehmer's (1983) study, for example, participants had to predict the level of the fictitious disease 'Morbus Brugonit' from the amount of 'Raviance' and 'Victine' in patients' bodies. Ashton's (1992) and Johnson and Kaplan's (1991) participants, who were professional auditors, had to predict bond ratings from three financial ratios and the risk for products in a company's catalogue becoming obsolete from five cues, respectively. Both Hagafors and Brehmer (1983) and Ashton (1992) found that external accountability increased the consistency with which participants applied their judgmental strategy (measured as the multiple correlation between the cues and participants' predictions). In Johnson and Kaplan's (1991) study, however, no significant difference between accountable and non-accountable participants was observed. Ashton (1992) and Johnson and Kaplan (1991) also reported a higher consensus between accountable participants than non-accountable participants, as indicated by the increased mean correlation between the predictions of pairs of participants. Finally, external accountability has been found to result in an increased self-insight of decision makers into their own judgement and decision processes. This is indicated by an increased correlation between beta regression weights and self-reported cue weights for accountable compared to non-accountable participants (Cvetkovich, 1978; Johnson & Kaplan, 1991).

If objective criteria exist, which allow a comparison of participants' predictions with actual values, participants' accuracy of judgement may be assessed. This was the case in Ashton's (1992) study, which found that accountable participants made more accurate judgements than participants who had not been made accountable. Studies employing other paradigms also have reported a higher accuracy of judgements under accountability pressure. In a study by Mero and Motowidlo (1995), for example, one group of participants was held accountable for their ratings of the performance of four simulated subordinates displayed on video, while another group was not held accountable. Accountable participants rated the displayed performance more accurately than non-accountable participants, as indicated by

Cronbach's (1955) index of differential accuracy, which measures how well raters can differentiate between different levels of performance on different dimensions. Although the primary focus here are judgement and decision processes, it should be mentioned that accountability has been found to improve other types of performance as well. For example, when asked to tear as many pages out of a phone book as possible within a certain time interval, accountable Type A individuals showed a better performance than non-accountable Type A individuals. For Type B individuals, no significant difference between accountable and non-accountable participants was observed (Yarnold, Mueser, & Lyons, 1988). This result suggests that there are inter-individual differences with regard to susceptibility to accountability manipulations. Other personality variables that have been found to increase people's response to external accountability include high self-monitoring, high social anxiety, a low individuating tendency, and low dogmatism (Chen, Shecter, & Chaiken, 1994; Fandt & Ferris, 1990; Lerner, 1994; Tetlock, Skitka, & Boettger, 1989).

External accountability has also been found to enhance people's memory performance. Chaiken (1980, exp. 1), for example, observed that participants recalled more arguments of a persuasive message if they had been told that they would be interviewed on the same topic at a second session than if they had been told that they would be interviewed on a different topic.

Finally, a number of studies have demonstrated that external accountability may have a de-biasing effect in those cases where a more thorough information processing results in a normatively correct response (Simonson & Nye, 1992, see also Lerner & Tetlock, 1999, for an extensive review). Externally accountable participants have repeatedly been shown to be less susceptible to a primacy effect in information processing (Freund, Kruglanski, & Shpitzajzen, 1985, exp. 3; Kruglanski & Freund, 1983, exp. 1; Tetlock, 1983b; Schadewald & Limberg, 1992; Kennedy, 1993; Webster, Richter, & Kruglanski; 1996). Wells, Petty, Harkins, Kagehiro, and Harvey (1977), Tetlock (1985b), and Lerner, Goldsmith, and Tetlock (1998) have demonstrated that accountability eliminates the fundamental attribution error, that is, the tendency to attribute behaviour to personal traits rather than situational factors (Ross, 1977). External accountability has also been found to reduce overconfidence in a personality prediction task (Tetlock & Kim, 1987), when giving eyewitness testimony (Kassin, Rigby, & Castillo, 1991), and when judging one's driving skills (McKenna & Myers,

1997). In addition, it seems to reduce bias in people's co-variation judgements. Slugoski, in an unpublished study, found that external accountability attenuated the illusory correlation effect, that is, the tendency to perceive a correlation between two classes of events when no such correlation is present in the data (Hamilton & Gifford, 1976). Under certain circumstances, external accountability may decrease people's reliance on stereotypes, for example, when they are not under time pressure (Kruglanski & Freund, 1983, exp. 2). Participants in Boudreau, Baron, and Oliver's (1992) study used fewer trait terms and described more specific behaviours when anticipating communication with an expert, and Foschi (1996) found that external accountability reduced a typical bias in the evaluation of task competence, namely that women are held to a stricter standard of competence than men. External accountability can also help to eliminate the tendency of incidental affect to elicit stereotypical information processing. Bodenhausen, Kramer, and Süsser (1994), for example, found that participants who had been put into a happy mood relied more on stereotypes than participants in a neutral mood, unless they had been made externally accountable. Furthermore, a number of biases identified by the Kahneman and Tversky research programme have been found to be attenuated under external accountability. Kruglanski and Freund (1983, exp. 3) showed a diminished anchoring and adjustment bias when participants had been made externally accountable, and Takemura (1993) demonstrated that participants who had to justify their decision did not show a framing effect but were more likely to choose the risky option in *both* gain and loss frames. A number of authors (e.g., Brockner, Shaw, & Rubin, 1979; Conlon & Wolf, 1980; Kirby & Davis, 1998; McLean Parks & Conlon, 1990; Simonson & Nye, 1992; Simonson & Staw, 1992) have provided evidence for the fact that when people are externally accountable, they are less likely to show escalation of commitment to a failing course of action and more likely to ignore sunk costs. This is in contrast to the effect of internal accountability on sunk costs reported earlier and provides further evidence for the suggestion that external accountability makes different behavioural standards salient than internal accountability. Finally, in the applied context of pilot decision making in cockpits, Mosier, Skitka, Heers, and Burdick (1998) could show that external accountability decreases the so-called automation bias. This term refers to omission and commission errors resulting from over-reliance on cues provided by automated aids, for example, not taking appropriate action because of not being informed of an imminent problem by the automated aid, or immediately acting upon automated information when there is other information in the environment that

contradicts or is inconsistent with the automated cue. This finding and the results of the studies reported above all strongly suggest that external accountability results in a more thorough processing of all available information.

Interestingly, external accountability has been found to reduce biased judgement and decision processes not only in individual but also in group decision making. Kroon and co-workers (Kroon, 't Hart, & van Kreveld, 1991; Kroon, van Kreveld, & Rabbie, 1991, 1992) have provided evidence for the fact that external accountability diminishes people's susceptibility to the so-called groupthink phenomenon, where the desire to reach a unanimous group decision overrides a motivation to critically analyse decision alternatives, resulting in a number of symptoms of defective decision making (Janis, 1982). In particular, external accountability inhibited the occurrence of pressure towards unanimity, encouraged all group members to participate actively in the discussion (thereby counteracting one of the antecedents of groupthink, promotional leadership), and resulted in less risky decisions. These effects were stronger when participants were made individually accountable (i.e., when they were told that they would each have to account for the group's decision and their own contribution to the discussion) than when they were made collectively accountable (i.e., when they were told that they would have to account for their group's decision together), although collective accountability was still more beneficial than no accountability (Kroon, 't Hart, & van Kreveld, 1991). Also, men seemed to be more affected than women (Kroon, van Kreveld, & Rabbie, 1992). Fandt (1991) found that making team members both individually and collectively accountable resulted in more interdependent behaviour, that is, more communication and co-ordination between group members in their effort to accomplish their task, which, in turn, lead to higher team satisfaction and higher perceived team success than no accountability. Unfortunately, no objective measure of team success was taken in this study, so that it cannot be established whether the increased subjective perception of success had any objective correlate.

The previous discussion of studies on judgement and decision biases seems to suggest that external accountability invariably reduces or eliminates such biases. This is, however, not the case. There are a number of biases that are either not affected or even negatively affected by external accountability. Which these are and under what circumstances external accountability will have a beneficial, no, or a detrimental effect will be discussed in more

detail later in this chapter. One reason for an impairment which is relevant to the discussion here, is that in certain situations the induced tendency to consider as much information as possible may backfire. A prime example is the dilution effect, the tendency for people to inappropriately take non-diagnostic information into account, which dilutes the extremity of predictions people would arrive at if they only used diagnostic information (Nisbett, Zukier, & Lemley, 1981; Zukier, 1982). Both Tetlock and Boettger (1989) and Tetlock, Lerner, and Boettger (1996) have demonstrated that external accountability increases participants' tendency to exhibit a dilution effect. Similarly, Stewart, Billings, and Stasser (1998) found that external accountability not only failed to enhance participants' tendency to share unique information with other members in their discussion group but even decreased it relative to a 'no accountability' condition, because of an increased focus on irrelevant details. At the same time, accountable participants were no more likely than non-accountable ones to identify the correct solution.

Whereas external accountability in many cases seems to have a beneficial effect, by inducing a deeper and more effortful information processing, the findings on the dilution effect and information sharing in groups suggest that external accountability can also impair judgement and decision processes. We will now turn to studies that have provided evidence for this impairment by suggesting that external accountability increases the individual's defensiveness.

Increased defensiveness

Commonly reported negative effects of external accountability point to an increased defensiveness on the part of the decision maker and include the following indicators:

- decisions that are expected to appeal to the evaluator's preferences and/or norms (Adelberg & Batson, 1978; Antonioni, 1994; Brief, Dukerich, & Doran, 1991; Brockner, Rubin, & Lang, 1981; Chinburapa, Larson, Brucks, Draugalis, Bootman, & Puto, 1993; Klimoski & Inks, 1990; Pennington & Schlenker, 1999; Schoenrade, Batson, Brandt, & Loud, Jr., 1986; Tetlock, 1983a; Tetlock, Skitka, & Boettger, 1989)

- decisions that are based on more generally accepted reasons that do *not* depend on knowledge of the particular evaluator's preferences (Aaker, 1991; Bowen & Qiu, 1992; Curley, Yates & Abrams, 1986; Simonson, 1989; Taylor, 1995)

- a focus on negative consequences and risk avoidance (Ganzach, 1993, exp. 3; Gibbins, 1984; Lord, 1992; McLean Parks & Conlon, 1990; Tetlock & Boettger, 1994)

- a defensive information search (Doney & Armstrong, 1996; Emby & Gibbins, 1988; Gibbins & Emby, 1985; Fandt & Ferris, 1990; Sanbonmatsu, Akimoto, & Biggs, 1993)

- an enhanced competitiveness in negotiators (Bass, 1966; Benton, 1972; Ben-Yoav & Pruitt, 1984; Carnevale, Pruitt, & Seilheimer, 1981; Druckman, 1968; Haccoun & Klimoski, 1975; Klimoski & Ash, 1974; Pruitt, Kimmel, Britton, Carnevale, Magenau, Peragallo, & Engram, 1978; Roloff & Campion, 1987)

- decision avoidance (Roberts, Stout, & Halpern, 1994; Staw & Boettger, 1990; Tetlock & Boettger, 1994)

- an enhanced tendency to rely on stereotypical information processing (Gordon, Rozelle, & Baxter, 1988, 1989; Hattrup & Ford, 1995, exp. 2)

Several studies have suggested that under external accountability, individuals are more strongly driven by a concern to avoid controversy and criticism than they are by the wish to reach an optimal conclusion, resulting in decisions that are easy to justify. One feature that makes decisions easy to justify is when they match the evaluator's preferences. Adelberg and Batson (1978) were among the first to demonstrate that the tendency to choose according to the preferences of one's evaluator(s) when having been made externally accountable can have considerable negative consequences. In their study, they told their student participants about a fictitious agency whose purpose was to provide financial aid for students who were in danger of having to drop out of university because of financial problems. Participants were told that they took part in a pilot study which tested whether students could function as local administrators of this programme. Each participant was given a certain budget and had to decide on six applicants with different personal histories and different levels of financial need. For each applicant, a minimal survival estimate was provided which was said to be necessary

in order to prevent the applicant from having to leave and, therefore, for the financial support to be effective. Participants were either not made accountable (i.e., were told that their decisions would remain anonymous), were made accountable to the service providers (i.e., were told that they would meet with an employee of the agency who would review their decisions), or were made accountable to the recipients of the service (i.e., were told that they would meet individually with each applicant to inform them about the decision). Also, the budget participants were given was either sufficient or insufficient to meet all applicants' needs. The main dependent variable was the amount of wasted resources, that is, the amount in dollars awarded below the specified minimum survival estimate. As predicted, Adelberg and Batson found that participants in the inadequate-resources conditions wasted more money than participants in the adequate-resources conditions, and within the inadequate-resources conditions those participants who were accountable to either the providers or the recipients of the service wasted more money than participants who were not accountable. This was because participants in these conditions were more concerned with pleasing the people they were accountable to than with providing effective help, as indicated by an analysis of the strategies participants applied to distribute their inadequate budget. Non-accountable participants mainly used an 'all to some' strategy which involved awarding substantial amounts to some applicants and nothing to others. This strategy minimised waste of resources. Participants in the 'recipient accountability' conditions, however, mainly applied a 'some to all' strategy, that is, they awarded all applicants some money. Since the amount awarded frequently was less than the minimum survival estimate, waste in this condition was high. Finally, although participants in the 'provider accountability' conditions, like those in the 'no accountability' conditions, primarily adopted an 'all to some' strategy, their waste was highest overall, because the amount they gave to their selected applicants often was less than was needed. The data therefore suggested that in the 'provider accountability' conditions, the salient norm was to spend as little as possible, whereas in the 'recipient accountability' conditions a fairness norm was salient, which prescribed to distribute resources equally among applicants. Importantly, when resources were insufficient (arguably, a common state of affairs for social services) external accountability, no matter whether to providers or recipients, resulted in a less effective use of resources than no accountability.

Similarly, a study by Pennington and Schlenker (1999) showed that when participants, who believed that they were real judges in a cheating case under adjudication by a student honour court, expected to discuss their decision with the accused student, they were more lenient in the punishment they recommended and wrote justifications that were more sympathetic to the student than when they expected to have to explain their decision to the professor who had brought forward the accusations, or to an official from the student honour court. These effects were still observed, although to a lesser extent, when participants made their decisions after being informed that the anticipated meeting had been cancelled, suggesting that external accountability not only resulted in normative influence but also prompted participants to genuinely change their perspective, consistent with informational influence. However, before such a conclusion can be drawn, it is important to rule out that participants may have believed that a meeting would take place at a later point.

Other evidence for the fact that under external accountability, individuals try to match their behaviour closely to a salient norm in order to gain the approval of the person(s) they are accountable to (or, as Bazerman, Tenbrunsel, and Wade-Benzoni, 1998, have put it, show a preference for 'should' over 'want' options), includes Schoenrade, Batson, Brandt, and Loud's study (1986). This demonstrated that female undergraduates, given a choice between benefiting themselves and benefiting another female undergraduate, were more likely to benefit the other person when they anticipated accountability to this person. Klimoski and Inks (1990) found that participants in the role of supervisors anticipating face-to-face feedback sharing with a poorly performing subordinate rated the subordinate's performance significantly more positively than participants who either expected to give written feedback or not to share any feedback. This effect has also been observed for upward appraisal ratings (Antonioni, 1994), where subordinates who were accountable for their ratings of their supervisor's performance made significantly more positive judgements than those who gave anonymous ratings. Finally, Brockner, Rubin, and Lang (1981) found that participants adopted their audience's view on what would be the best investment strategy to win them money in a jackpot. When the audience was expected to be cautious, participants made conservative investment decisions, when it was expected to be risky, participants made risky investment decisions.

The previous discussion has suggested that one way of coping with external accountability pressures is making a decision that is assumed to please the audience one is accountable to. However, a choice may also be perceived as easy to justify, and therefore preferred under external accountability, when it is *not* dependent on knowledge of the particular evaluator's preferences but is based on more generally accepted decision principles instead. One such principle, according to Curley, Yates, and Abrams (1986), may be ambiguity avoidance. In their studies, participants showed ambiguity avoidance when choosing between two lotteries with identical outcomes, one of which was ambiguous because the probability of the outcomes was not known. Curley et al. argued that the preference for the unambiguous lottery could be explained by participants' belief that a choice that avoided ambiguity was most socially acceptable. This explanation was borne out of the fact that the anticipation of being evaluated by others was the only one out of six manipulations that systematically affected participants' choice behaviour in this situation. A possible reason for why ambiguity avoidance may be perceived as more justifiable, Curley et al. suggested, could be that external accountability focuses participants' attention on possible negative decision outcomes, and participants choose the non-ambiguous option in order to avoid the outcome of losing with an ambiguous option that turns out to have a worse probability of success than the unambiguous one. This assumption received some support by Bowen and Qiu's (1995) study, which showed that, when faced with a choice between two unattractive alternatives, one of which was unattractive because it was ambiguous and the other because it had a low probability of success, participants were willing to pay to buy information that reduced uncertainty about the ambiguous option, whereas they were not willing to do so when the unambiguous alternative was attractive, thereby having a good reason already to choose it over the ambiguous option.

Another example of more justifiable decisions under external accountability is Simonson's (1989) demonstration that both the attraction and compromise effect are enhanced when participants expect to have to justify their decision. The attraction (or asymmetric dominance) effect, first described by Huber, Payne, and Puto (1982), refers to 'the ability of an asymmetrically dominated or relatively inferior alternative, when added to a set, to increase the attractiveness and choice probability of the dominating alternative' (Simonson, 1989, p. 158). For example, if participants have to choose between two brands of beer, one

cheaper and the other of better quality, when a third brand is added to the choice set which is dominated by one of the alternatives (e.g., because of being more expensive and no better quality), the choice ratio (i.e., market share) of the dominating alternative goes up compared to a situation where the dominated alternative is absent from the set. This effect is enhanced when participants anticipate having to justify their decision (see also Aaker, 1991). According to Simonson (1989), this is because the asymmetric dominance relationship provides the only reason for a choice that does not depend on particular attribute weights or trade-offs and therefore is independent of any evaluator's particular preferences. A similar argument may be made for the amplification of the compromise effect under external accountability, that is, the tendency for an alternative to become more popular when it becomes a compromise or middle option in a choice set. If the decision maker is unsure about what the person(s) he or she is accountable to prefer, then a compromise alternative seems to be the safest choice with the smallest maximum error.

External accountability seems to focus people's attention particularly on the potential negative consequences of their decisions, arguably because the penalties for errors are more significant than the rewards for success (Lord, 1992). In Tetlock and Boettger's (1994) laboratory simulation of FDA (Food and Drug Administration) decisions about whether to allow drugs onto or take already approved drugs off the US pharmaceutical market, participants who anticipated to have to justify their decisions to an expert focused their thoughts more on negative aspects of the drug than non-accountable participants. Also, Ganzach (1993, exp. 3) found that when participants were asked to adopt the role of personnel manager in charge of hiring personnel for managerial positions rather than just to evaluate two job candidates, they used a conjunctive judgement strategy more often that relied more heavily on negative cues. Finally, McLean Parks and Conlon (1990) reported that decision makers who were blamed for a previously failed investment decision, thereby having accountability concerns made salient to them, focused more on negative cues when asked to state how much money they would continue to invest in the option that had failed.

Presumably as a consequence of focusing on the potential negative aspects of their decisions, individuals often show more conservatism and risk avoidance under external accountability. In Tetlock and Boettger's (1994) study, for example, accountable participants who had to decide on a medium or high-risk drug not yet on the market were particularly

reluctant to permit it. Lord (1992) showed that professional auditors who were accountable for their decisions to a senior partner of their firm, were less likely to issue an unqualified opinion than auditors who had been given a guarantee of anonymity.

Further indicators of the tendency to become more defensive under external accountability are confirmatory and symbolic information search. The latter refers to information search that is carried out for the primary purpose of justifying one's actions to others (O'Reilly, 1983). In a field survey by Doney and Armstrong (1996), purchasing managers indicated that both informal accountability (accountability to subordinates and/or peers) and official accountability (accountability to superiors) made them collect more information than they really needed, in order to show that they did a good job in case something went wrong later on, or for the purpose of making a good impression on others. A similar result was obtained by Fandt and Ferris (1990). In their study, customer service employees were asked to consider a scenario which, although hypothetical, was very close to their daily experience and involved making a decision and documenting this decision by writing a brief report and selecting information items from a provided list that would be contained in a brief oral report. Participants were either told that they had to stand in as a temporary supervisor for a day, with personal responsibility for any decisions made, and that their performance would affect their job evaluation and future promotion prospects ('high accountability' condition), or they were told that someone else had been made temporary supervisor and that they were not responsible for any decisions made nor would their job performance evaluation or future career prospects be affected ('low accountability' condition). In addition, decisions were either made under high ambiguity, where participants were told that a similar incident had occurred the previous week but they were not aware of their supervisor's decision then, or under low ambiguity, where participants were told how their supervisor had handled a similar incident the previous week. It was found that when accountability was high and ambiguity was low, participants used significantly more defensive information, that is, information that served to shift responsibility and cover mistakes, and mentioned significantly more positive aspects of their decision than in any other condition. Also, Sanbonmatsu, Akimoto, and Biggs (1993), in the context of a causal judgement task, showed that external accountability increased confirmatory information processing, that is, the tendency to mainly consider the likelihood of an outcome in the

presence of a target cause and to neglect the likelihood of the outcome in the absence of the target cause or in the presence of alternative causes, resulting in strong overestimations of the causality of a certain cue.

Defensiveness-enhancing effects of external accountability have also been observed in negotiation contexts. External accountability arises when negotiators act as representatives for a party and their constituents have the capacity to monitor their performance and to reward or punish them. In laboratory settings, the latter is usually manipulated by the degree to which each negotiator's earnings are determined by his or her team-mates. External accountability to their constituents has been found to increase negotiators' competitive behaviour, which can stand in the way of reaching the best possible negotiation outcome. Externally accountable negotiators are less likely to reach an integrative agreement which satisfies the needs of both parties (Benton, 1972; Carnevale, Pruitt, & Seilheimer, 1981), and the overall profit is lower than in negotiations between non-accountable representatives, usually because the profit of the less successful negotiator is lower (Pruitt, Kimmel, Britton, Carnevale, Magenau, Peragallo, & Engram, 1978; Ben-Yoav & Pruitt, 1984). In cases where the quality of solution is not determined by a quantitative criterion such as overall profit, but where a correct solution to a problem exists, such as the NASA moon survival task (Hall, 1971), externally accountable negotiators sometimes, but not always, produce solutions of a poorer quality (Haccoun & Klimoski, 1975). Externally accountable negotiators make more extreme initial offers (Roloff & Campion, 1987), are less willing to concede (Pruitt et al., 1978) and more often reject others' offers. They therefore usually take longer to reach an agreement, especially when groups formulate and adopt a position prior to negotiations, and are more likely to reach a deadlock (Bass, 1966; Druckman, 1968; Benton, 1972; Klimoski & Ash, 1974, Van Hiel & Schittekatte, 1998). Under external accountability, effective problem-solving behaviour, such as the exchange of information about values and priorities, is less likely to occur. Instead, so-called distributive behaviours can be observed more often, where negotiators try to dominate the other party by using threats, positional commitments, arguments aimed at the other to concede and attempts to raise one's status (Pruitt et al., 1978; Carnevale, Pruitt, & Seilheimer, 1981; Ben-Yoav & Pruitt, 1984). A definition of the negotiation situation as highly competitive is reflected in externally accountable negotiators' perceptions of themselves, their constituents and the other party. In Roloff and Campion's

(1987) study, externally accountable negotiators thought they were perceived as less co-operative and perceived the other party as less co-operative than unaccountable negotiators. Benton (1972) reported that externally accountable negotiators perceived their constituency as more competitive and more often thought their team-mates expected them to win than negotiators who were not accountable. Finally, under external accountability negotiators typically report more pressure, frustration and difficulty than without accountability (e.g., Klimoski & Ash, 1974). All these findings suggest that in negotiation contexts, accountability makes a competitiveness norm salient, where negotiators are meant to try to achieve a maximum outcome for their constituents and convey an image of toughness. The typically strong concern for negotiators' own and lack of concern for the other party means that both the negotiation process and outcome will be sub-optimal and that accountability will have a detrimental effect. However, when negotiators believe that they are rewarded for their objectivity rather than partisanship, accountable negotiators are better able to perceive interests compatible with the other party than non-accountable negotiators (Thompson, 1995).

The enhanced pressure individuals seem to perceive when they are made accountable often results in decision avoidance tactics, as proposed in Janis and Mann's (1977) conflict model of decision making. In Tetlock and Boettger's (1994) FDA simulation study, participants who anticipated having to justify their decisions to an expert and who were confronted by a medium or high-risk drug not yet on the market, were particularly likely to show procrastination or buck-passing, that is, were particularly likely to advise deferring the decision for another year or passing it on to another decision making body, despite the fact that it was extremely unlikely that new evidence would emerge during this time period or that the other agency would be able to make a more informed decision. Roberts, Stout, and Halpern's (1994) analysis of decision making in a military context also suggested that soldiers at lower hierarchy levels typically relieve themselves of unpleasant feelings of responsibility by reporting problems to decision makers at the next higher level. This is consistent with Staw and Boettger's (1990) finding that external accountability impedes task revision in subordinates because they pass responsibility on to a higher level.

Several authors have suggested that the stress induced by being accountable to others may also help to explain why participants who expect having to justify their judgements and decisions to others sometimes are more likely to rely on simple heuristics, for example,

stereotypes, than those who do not, contradicting research presented earlier that demonstrated a reduction of stereotypical information processing under external accountability. In Gordon, Rozelle, and Baxter's (1988, 1989) studies, for example, participants who rated simulated videotaped employment interviews of female job applicants, rated younger applicants (25 years old) more positively, older applicants (40 and 55 years old) more negatively, and made more dispositional trait attributions when they expected to discuss their impressions of applicants with personnel directors compared to when they thought their responses would remain confidential. Similarly, Hattrup and Ford (1995, exp. 2) found that when participants were asked to search information about several targets prior to judging their suitability for membership on a committee, they searched significantly less information for targets with occupational labels when they had been made accountable than when they had not been made accountable. This supports the notion that external accountability may induce high levels of stress, which, in turn, results in more simplistic thinking (Janis & Mann, 1977).

In conclusion, both internal and external accountability have been shown to affect judgement and decision processes in two seemingly contradictory ways. On the one hand, they induce deeper and more complex information processing, on the other hand, they make the decision maker more defensive and enhance his or her need to make a decision that can be easily justified rather than an optimal one. In order to explain why and under what circumstances internal and external accountability have these different effects on decision processes, three classifications of accountability seem of particular relevance. These are (1) accountability to an audience with known vs. unknown views, (2) pre-decisional vs. post-decisional accountability, and (3) procedural vs. outcome accountability. In the following, these classifications will be introduced and the effects of each type on decision processes and outcomes discussed.

Accountability to an audience with known vs. unknown views

The first classification that allows predicting when external accountability will increase the depth and complexity of information processing and when it will increase a decision maker's defensiveness instead, is that of accountability to an audience with known vs. unknown views. Tetlock (1985a, 1991, 1992), in his social contingency model of

judgement and choice, has suggested that external accountability will result in more thorough information processing only when decision makers are accountable to an audience with unknown views. Otherwise, because people are 'cognitive misers' and prefer to expend minimal effort, they will just adopt the position which is most likely to be accepted by those to whom they feel accountable. Hence, decision makers' knowledge or assumption about the views of the person(s) they have to justify their decisions to moderates the effects of external accountability on decision processes. If they know or at least have certain expectations about the views of the person(s) they are accountable to, decision makers will make a decision that is consistent with these views and will not put increased cognitive effort into the decision process. Tetlock calls this an *acceptability heuristic*. Only when the most socially acceptable option is not obvious, does external accountability result in vigilant information processing in order to identify the alternative that is most defensible. The latter process Tetlock refers to as *pre-emptive self-criticism*. Individuals, given a situation of normative ambiguity and pressure to justify their decisions, will consider a variety of information, in order to anticipate arguments that critics could raise against their positions. This, according to Tetlock (1992), may be viewed as an adaptive strategy to protect both one's self-image and social image.

We have already referred to evidence which supports Tetlock's hypothesis, for example, Tetlock (1983a), who found that the integrative complexity of participants' thoughts was lowest when they had been made accountable to someone with known views (either liberal or conservative) and highest when they had been made accountable to someone with unknown views, with no accountability in-between. Also, Mero and Motowidlo (1995) showed that the accuracy of participants' performance ratings was only improved if they had no expectations regarding the desired outcome of their ratings. When they had been told beforehand that their previous ratings had been too low, participants systematically overestimated the performance of others.

When individuals are not sure about the standards and norms of the audience they are accountable to, they often attempt to guess the views of their audience and behave accordingly. Given that individuals are prone to a false consensus effect (Ross, Greene, & House, 1977), where they overestimate the extent to which others agree with their opinion, it is not surprising that this guess is often biased towards the individual's own view and/or preferences. Weigold and Schlenker (1991, exp. 1), for example, found that low risk takers

became more risk-averse when accountable, whereas high risk takers became slightly, although not significantly, more risk-taking. A second study confirmed that high risk takers admired and thought that also other people admired risk takers more than risk avoiders. It is also likely that the reason why participants in Takemura's (1993) study who had to justify their choice in a monetary decision task chose the risky option more often than those who were not required to justify their choice, was that the former assumed that risk taking was more valued. The same seems to have been true in Kirby and Davis's (1998) study. Participants whose decisions were monitored allocated smaller funds to risky strategies than participants whose actions were not monitored.

A number of studies have directly investigated the effect of explicitness of audience norms on individuals' tendency to adopt these norms. In Brief, Dukerich, and Doran's (1991, exp. 2 and 3) study, for example, participants were asked to role-play a member of a board of directors of a pharmaceutical company and had to make a decision on how to react to plans to ban a dangerous drug marketed by their company. They were made accountable to the chairman of the board of directors and either told explicitly how the chairman intended to resolve the case, or it was suggested to them that the chairman had either Smithian or humanitarian views, by describing his stand on other issues. As expected, participants made choices that were more consistent with the chairman's value system when they explicitly knew how he would choose compared to when they merely knew about his values. In Pennington and Schlenker's (1999, exp. 2) study reported above, only when the preferences of the central parties were clearly stated did the nature of the audience with whom participants expected to meet and discuss their decisions have a significant impact. Also, Fandt and Ferris's (1990) previously reported study suggested that participants only engaged in defensive information search when they had been told how the person they had been made accountable to had handled a similar incident before.

Research in the area of conformity (e.g., Hollander, 1960) suggests that individuals should be more likely to deviate from the norms they perceive their audience to endorse and not adapt their behaviour to these norms, when they themselves have high status. In negotiation studies, accountability effects have indeed been found to depend on whether negotiators have high status, for example, because of being elected (Klimoski & Ash, 1974) or because the nature of their mandate is said to be based on competence rather than liking

(Haccoun & Klimoski, 1975). In Haccoun and Klimoski's (1975) study, negotiators perceived more flexibility in their roles and deviated to a greater extent from their team's position when they had been elected on the basis of their performance in a previous task, particularly when they were evaluated by strangers, whereas they were found to be very loyal to their groups and did not deviate from their team's position when they had been elected on the basis of likeability[3]. This suggests that individuals are more likely to adhere to their evaluators' norms when they care for their affection, and, therefore, knowledge of the particular relationship between an agent or spokesperson and his or her constituency is a very important factor for correctly predicting accountability effects. Similarly, one would expect that the stronger an evaluator's reward or punitive power, the stronger accountability effects should be. The first assumption was supported in a study by Cvetkovich (1978), in which participants who were accountable to a friend recalled their judgement strategy with greater accuracy than participants who were accountable to a stranger or not accountable at all. The second hypothesis was supported in a study by Gordon and Stuecher (1992), in which students who expected to have to justify their evaluations of a teacher to a faculty member wrote more grammatically complex evaluations than students who expected to have to justify their evaluations to another student.

In some cases, the observed effects of external accountability may be inconsistent, because the individual feels accountable to more than one party. In a medical task context, Chinburapa, Larson, Brucks, Draugalis, Bootman, and Puto (1993) demonstrated that physicians who were told that their choice of a hypothetical anti-infective drug would be reviewed and evaluated by an infectious disease specialist, paid more attention to the costs of the different drugs and more often rated the cost attribute as the most important, but at the same time were more likely to choose the drug with the highest cost (but also highest cure rate) compared to participants who had not been made accountable[4]. This suggests that accountability can be a dynamic process; during information search, accountability to their peers seems to have been salient to the participants, but when it came to choosing a drug,

[3] In the latter condition, negotiators also took a longer time to reach agreement when they were evaluated by their own team compared to strangers, suggesting that the effects of external accountability to one's own constituency are different from pure evaluation apprehension effects.

[4] This effect was only observed in a high complexity condition.

accountability to the patient must have been more salient. We have already referred to the problem of multiple accountabilities in Chapter 1. When individuals are accountable to different people with conflicting goals, it becomes very difficult indeed to be able to correctly predict accountability effects. In negotiation contexts, role conflicts based on accountability to both negotiators' own constituents and the other party have been found to have a positive impact. Pruitt and Carnevale's (1982) dual-concern model of negotiation proposes that representatives who are concerned about a satisfactory outcome both for their own group and the other group are most likely to adopt a constructive problem solving strategy. Accordingly, Carnevale and Mack (reported in Carnevale, 1985) found that participants who had to justify their opinions on several issues to two other people with contrasting attitudes wrote more complex statements of their opinions than participants who were not told anything about the others' attitudes or expected them both to have the same attitude.

While individuals' concern about maintaining a good relationship to other people clearly seems to have an effect on their behaviour, behaviour is obviously also a function of individuals' personal goals (Ben Yoav & Pruitt, 1984). Hence, the congruency between the individuals' own norms and values and those of others they feel accountable to should moderate effects of external accountability. When individuals' own norms coincide with those of the people they feel accountable to, conformity should be stronger than when they conflict. This issue has been given little attention so far in the literature. In a field study conducted at two departments of the Dutch Riot Police Academy by Kroon, van Kreveld, and Rabbie (1991), incongruence between an anti-violence norm introduced by superiors and the norm already existing within the institution (and hence, presumably, the individual members of the institution) resulted in reactance, where participants who had been made accountable declined the norms introduced by the authorities and showed reluctance to co-operate. This last finding underlines the importance of legitimacy of accountability. Accountability to an audience that is not perceived as legitimate is likely not to have any beneficial effects and may even backfire (Lerner & Tetlock, 1999). It also becomes clear that, unless certain norms are explicitly made salient to the participant, accountability effects are difficult to predict and often involve hindsight guessing about which norms the individual may have tried to conform to.

The distinction between accountability to an audience with known vs. unknown views also sheds some light on the question under what circumstances external accountability may have a beneficial, no, or detrimental effect on cognitive biases. Simonson and Nye (1992) have suggested that external accountability only has a de-biasing effect, if (1) decision makers can anticipate which response will be rational and will be evaluated most favourably, and they are less likely to choose this response when not being accountable, or (2) the normatively correct decision can be identified by the more thorough and complex information processing instigated by accountability to an audience with unknown views. For example, in Simonson and Nye's (1992) studies, accountable Business Administration students, who were aware of the normative rule that sunk costs should be ignored, were less susceptible to the sunk cost effect than non-accountable participants. They were also more likely to produce rational as opposed to emotional reasons for their choice. Furthermore, external accountability failed to eliminate biases in cases where the normatively correct response was unlikely to be identified through more effortful information processing, namely, the tendency to ignore sample size and base rates, the conjunction fallacy, choice-matching preference reversals (see also Selart, 1996) and overconfidence associated with judging the likelihood of events that are causally related. In the case of these latter biases, individuals do not seem to be aware of what would constitute a rational response nor does a more thorough information processing help them to arrive at the normatively correct response. On the contrary, in the case of the dilution effect, as described previously, a more thorough information processing makes matters worse, because the individual pays too much attention to irrelevant details. That the amplification of the dilution effect under external accountability is also partly an effect of the individual's enhanced tendency to conform to salient norms, has been shown by Tetlock, Lerner, and Boettger (1996). In their study, accountable participants did not exhibit the dilution effect when Gricean norms of conversation, such as truthfulness and relevance of the information provided, were explicitly deactivated, whereas they did show a dilution effect when conversational norms were explicitly primed as well as in a 'no prime' control condition. Non-accountable participants showed the effect across norm activation conditions and most strongly under activation of conversational norms. The tendency to adhere more strongly to salient norms when participants expect to have to justify their judgements and/or decisions to others, also explains why biases such as the attraction effect, the compromise effect, the status quo effect or stereotypes become stronger under external accountability. These biases are

expression of the individual's need to make decisions that are easy to justify, because they are expected to appeal to the evaluator.

Pre-decisional vs. post-decisional accountability

A second important distinction that allows explaining the differential effects of accountability is that of pre-decisional vs. post-decisional accountability. As the term suggests, pre-decisional accountability is present if decision makers know before they embark upon a decision that they will be called to account for it afterwards. Post-decisional accountability occurs when decision makers are asked to account for their decision after it has been made, and often also only after its outcome has become known. Typically, pre-decisional accountability results in pre-emptive self-criticism, indicated by a thorough, complex information processing, where the individual tries to anticipate the objections of potential critics, whereas post-decisional accountability results in *defensive bolstering* of the decision, where people try to generate as many reasons as they can for why they are right, accompanied by less complex, rigid, and evaluatively more consistent thought processes (Tetlock, 1983, 1985a, 1992).

Supportive evidence for this hypothesis was obtained in a study by Tetlock, Skitka, and Boettger (1989). Participants were asked to list their thoughts about each of four controversial issues, under conditions of either no expectation of having to justify their attitude, an expectation of having to justify their attitude before they were indicating it, or an expectation of having to justify their attitude only after they had been indicating it. In addition, participants in the accountability conditions expected to have to justify their decision to either a person with unknown, with liberal, or with conservative views. As predicted, the integrative complexity of participants' thoughts was highest when they had been made accountable to an audience with unknown views before they had committed themselves to a particular stand. It was also higher in the pre- compared to the post-commitment conditions. This was not true, however, when participants had not been made accountable, suggesting that the timing of the accountability manipulation alone was not sufficient to produce simplification of thought. Also, in line with predictions, it was found that participants shifted

their view to that of their prospective audience only when they knew the views of the audience, and when they had not committed themselves previously.

Further extensive support for the assumption that post-decisional accountability enhances defensiveness comes from the numerous studies on the effects of cognitive dissonance mentioned earlier, which have demonstrated that once participants have committed themselves, for example, by choosing an alternative, they search information and evaluate it in such a way that their previous choice is supported. This was also true in the studies of escalation of commitment reported earlier, where retrospective personal responsibility for a failed course of action resulted in more resources being allocated to this course of action. Interestingly, Kirby and Davis (1998) were able to show that making participants believe that their principal would monitor their decisions, could offset the negative effects of retrospective personal responsibility to some extent. They found significant main effects for both personal responsibility and external monitoring. Whereas personally responsible participants allocated more funds to a failing strategy than non-responsible participants, participants whose actions were monitored made smaller allocations than participants whose decisions were not monitored. This is a nice illustration of the fact that internal and external accountability can have very different effects, depending on the norm that becomes salient. In this study, personal responsibility for a previous failure seems to have strengthened the individual's wish to enhance their self-esteem and prove to themselves that their previous decision was a good one, resulting in even more funds allocated, whereas external monitoring seems to have made the principal's interests salient, thereby diminishing the tendency to escalate commitment.

Procedural vs. outcome accountability

A final important distinction between different types of accountability has been suggested by Simonson and Staw (1992, see also Peecher & Kleinmuntz, 1991, and Siegel-Jacobs & Yates, 1996). They contrast process or procedural accountability with outcome accountability. Procedural accountability is accountability for the procedure the decision maker uses to arrive at his or her decision. An evaluation of the decision is based exclusively on the quality of the procedure used. An example may be a situation in which an auditor has

to justify an audit procedure to a senior partner of his or her firm. Outcome accountability, on the other hand, means that the decision maker is accountable for the outcome of his or her decision and is evaluated on the basis of this outcome. For example, a manager may be accountable for the results of his or her investment decisions. It should be noted, however, that in practice this difference may often be blurred, since accountability often only becomes salient when an unexpected negative outcome occurs. This negative outcome is likely to enhance outcome accountability, even though the individual may be subject to procedural accountability only. For example, a medical doctor may feel accountable for adverse effects of a treatment he or she has recommended (outcome accountability), although he or she would normally only be accountable for how a particular course of treatment is chosen (procedural accountability). Siegel-Jacobs and Yates (1996) have pointed out two important differences between procedural and outcome accountability, in terms of (1) guidance they offer the decision maker of how to achieve a positive evaluation, and (2) perceived ease of improving performance. Outcome accountability puts pressure on the decision maker to produce a successful outcome, but does not suggest how this may be achieved. Procedural accountability, on the other hand, signals to the decision maker that extra effort put into the decision process will protect him or her from a negative evaluation. A decision maker under procedural accountability therefore will find it relatively easy to improve his or her performance, since, as long as sufficient effort is expended or a certain previously specified procedure is followed, the decision maker will be exonerated from any blame for a negative outcome. Outcome accountability, however, induces a high level of stress, because a negative outcome can occur despite following a correct procedure or expending a lot of effort. These differences between outcome and procedural accountability are reflected in the differential effects they have on decision processes. Generally, procedural accountability has found to be beneficial, whereas studies employing outcome accountability have often shown detrimental effects.

For example, in Simonson and Staw's (1992) investigation of escalation of commitment, outcome accountability was manipulated by telling participants that if they made particularly effective or ineffective decisions, their data would be shared with other students and lecturers and their evaluation as a decision maker would be based on whether the course of action they initially recommended turned out to be most beneficial to their

company. Procedural accountability was manipulated by telling participants that if they used particularly effective or ineffective decision *strategies*, their data would be shared with other students and lecturers and their evaluation as a decision maker would be based on their use of strategies rather than the outcome of those strategies. As predicted, procedural accountability resulted in significantly lower subsequent allocations to a failed course of action than both a baseline condition which established personal responsibility for participants' initial decision and outcome accountability, thereby preventing escalation of commitment. Allocation of funds to the failing course of action under outcome accountability, although larger than in the baseline condition, was not significantly different from it. These results clearly suggest that procedural accountability reduces escalation of commitment, whereas outcome accountability does not.

In Siegel-Jacob and Yates's (1996) studies, participants had to predict the likelihood that each of a set of individuals held a particular attitude from a number of different cues. Participants were either not made accountable, were made accountable for their decision process (by being informed that there would be a post-experimental interview on why and how they used the information they did to arrive at their judgement), or were made accountable for their decision outcomes (by being told that they would receive feedback on their accuracy relative to the other participants in the study and that there would be a cash prize for the people with the top five scores). Procedural accountability improved calibration relative to no and outcome accountability and also resulted in more information being taken into account. Whether or not this had a beneficial effect on participants' performance depended on whether the set of cues the participants were provided with contained only valid predictors or a mixture of both valid and invalid predictors. When the set contained valid cues only, participants under procedural accountability showed improved discrimination (i.e., assigned high probabilities when the target event did indeed occur and low probabilities when it did not occur). When it contained both valid and invalid cues, however, discrimination was not enhanced, because irrelevant information was taken into account, that is, participants showed a dilution effect. Procedural accountability was also found to counteract the negative effects of outcome feedback on scatter, that is, significantly reduced the amount of variability or inconsistency in participants' judgements. Outcome accountability, on the other hand,

significantly increased the amount of scatter compared to no or procedural accountability and therefore resulted in lower accuracy overall.

In line with Siegel-Jacobs and Yates's (1996) results, Doney and Armstrong (1996) found that procedural but not outcome accountability had a positive effect on participants' information analysis, making them spend more time and effort on collecting and analysing information prior to a decision. Finally, Ordonez, Benson III, and Beach (1999) showed that making participants accountable for the strategy they used for screening a set of jobs resulted in more consistency with which the screening strategy was applied. Procedural accountability also magnified the difference between threshold values in a 'reject' condition, in which participants were told to eliminate any jobs they did not want to apply for, and an 'apply' condition, in which participants were asked to select any jobs they wanted to apply for. Participants' tendency to have a lower rejection threshold, that is, retain fewer jobs in the choice set if they were asked to select the jobs they did not want to apply for than if they were asked to select the jobs they did want to apply for, was stronger in the procedural accountability condition than in a 'no accountability' condition, once again suggesting that accountability intensifies individuals' adherence to salient norms.

This last point draws attention to an important qualification that has to be made when talking about the effects of procedural accountability. Although there is convincing evidence for the fact that the effects of procedural accountability are generally beneficial, it has to be kept in mind that whether or not this will be the case depends on whether the individual knows the correct or rational procedure for arriving at a judgement or decision. Siegel-Jacobs and Yates's (1996) results already suggested that individuals who expect to have to justify their judgement strategy may be prone to a dilution effect if the information set contains invalid predictors, because of their tendency to try to take *all* information into account. Also, some negotiation studies have suggested that when constituents are able to monitor the negotiation process rather than just being informed about the negotiation outcome, negotiators have a greater desire to appear strong, produce more competitive and contentious behaviour, and are less likely to reach an integrative agreement (e.g., Pruitt et al., 1978), presumably because they think that this is the kind of behaviour that would impress their constituents. Therefore, despite accountability for the negotiation process and not its outcome, a negative effect was obtained.

Summary and conclusions

This review of studies investigating the effects of internal and external accountability on judgement and decision making has identified two major, seemingly contradictory, trends. Both internal and external accountability, under some circumstances, increase the decision maker's depth of information search and complexity of information processing, resulting in less biased judgements and decisions. At the same time, however, internal and external accountability also seem to increase the decision maker's defensiveness and his or her need to make decisions that are easy to justify rather than optimal, indicated by an enhanced tendency to make decisions that appeal to the evaluator's preferences and rationalise previously made decisions. Three classifications of different types of accountability help to explain why sometimes one effect is more prominent and sometimes the other. These are accountability to an audience with known vs. unknown views, pre- vs. post-decisional accountability, and procedural vs. outcome accountability. Tetlock (1985a, 1991, 1992), in his social contingency model of judgement and choice, has proposed that when the decision maker is accountable to someone with unknown views and has not previously made a commitment, he or she will show pre-emptive self-criticism, that is, effortful, thorough, and complex information processing which attempts to anticipate the objections of potential critics. On the other hand, when the decision maker is accountable to someone with known (or anticipated) views and has not made any previous commitment, instead of engaging in more complex information processing, he or she will employ an acceptability heuristic that prescribes choosing the socially most acceptable option, without investing the additional effort of carefully analysing the other options. Finally, once the decision maker has committed him- or herself to a particular position and is made accountable only after this commitment, he or she will engage in defensive bolstering, a process of rationalising his or her previous choice. Another type of accountability that promotes defensiveness is outcome accountability, accountability that focuses the decision maker's attention on the outcome of his or her choice and may create high levels of stress, because the decision maker often does not have control over the decision outcome, even when the choice is carefully contemplated. Procedural accountability, on the other hand, where the decision maker is accountable for how he or she arrived at a decision rather than its outcome, is likely to enhance pre-emptive self-criticism and therefore to improve the decision process, unless the decision maker follows a sup-optimal procedure.

Based on the findings reported in this chapter, it is argued here that accountability is as a motivating agent which typically makes two different needs salient, a need to be accurate and a need to arrive at an easily defensible decision. Accountability is proposed to enhance the salience of *both* these needs. Whether or to what extent one need will dominate the other, will depend on the particular situation the individual finds him- or herself in. Accountability to someone with unknown views, pre-decisional accountability and procedural accountability all seem to enhance the individual's motivation to make an accurate decision, whereas accountability to someone with known views, post-decisional accountability, and outcome accountability all seem to enhance the individual's motivation to make a decision that will be readily accepted by his or her audience. Any observed accountability effects will therefore depend on the particular combination of these factors in the accountability manipulation.

Another distinction that seems to be important when trying to understand accountability effects is that between effects on the decision process and effects on the decision itself. Accountability clearly influences both *how* people think, that is, their decision process, and *what* they think, that is, the decision outcome or the particular preferences they express (Simonson & Nye, 1992). Whereas the beneficial effects of accountability seem to be mostly related to an improvement of the decision process, particularly a deeper information search and a more complex information integration, its detrimental effects, namely the tendency to become more defensive and make decisions that are easy to justify rather than optimal, are mainly observed in the particular decision outcome and are the result of a biased information evaluation. It may therefore be argued that an increase in the complexity of information processing and an increase in defensiveness are not really contradictory effects, but effects of accountability on different stages of the judgement and decision process, with the former mainly affecting information search and the latter mainly affecting information evaluation.

This suggests that decision makers engage in a process of self-regulation, using different cognitive mechanisms to try to achieve each of the different needs made salient by accountability pressures. In particular, decision makers are assumed to react to the enhanced need to be accurate by conducting a deeper information search and more complex information integration, whereas they react to the need to arrive at an easily justifiable decision with a biased information evaluation and integration. How exactly the decision maker may use these

different mechanisms to achieve their goals will be discussed in more detail in the following chapter.

A question that stills remains to be answered is whether, and if so, how the observed effects of accountability are related to its antecedent conditions, personal responsibility and evaluation apprehension. From the review of accountability studies above, it seems that personal responsibility as a precondition of internal accountability enhances the individual's motivation to make an accurate decision and thereby affects depth of information search and complexity of information integration, as there is no evidence to suggest that individuals become more defensive when they are held personally responsible for their judgements and decisions, unless this personal responsibility is retrospective. Evaluation apprehension, on the other hand, seems to enhance the individual's motivation to make a decision that can be easily justified, as it focuses the individual's attention on decision outcomes by making external norms salient, and thereby affects the extent of bias in information evaluation and integration. As mentioned above, the next chapter will analyse these processes in more detail and present a self-regulation model by Baumeister and Newman (1994) which provides a theoretical framework for them.

CHAPTER 4: MOTIVATED INFORMATION PROCESSING

In this chapter we will take a closer look at the processes of information search, evaluation and integration, and their susceptibility to different motivational goals of the decision maker. First, a general theoretical framework for motivated inference and decision processes proposed by Baumeister and Newman (1994) will be presented, which focuses on the effects of the need to reach an optimal conclusion (i.e., satisfy an accuracy goal) vs. the need to reach a particular conclusion (i.e., satisfy a directional goal). It will be followed by a review of the literature on various forms of confirmatory bias as the decision maker's preferred means to reach a desired conclusion. The discussion of evidence for confirmatory information search will largely focus on research conducted within the framework of dissonance theory (Festinger, 1957, 1964), whereas that for confirmatory evaluation and integration processes will refer to a theory that has spelt out in some detail where such biases may occur when multiattribute decisions are made, namely Svenson's (1992, 1996) differentiation and consolidation theory. This review of the motivated reasoning literature will provide a basis for the following chapter, in which a process model of accountability will be developed that aims to predict how accountability affects processes of information search, evaluation and integration in decision making.

Baumeister and Newman's (1994) framework of self-regulated inference and decision processes

A number of different authors, for example, Kruglanski (1980, 1989, 1990; Kruglanski & Ajzen, 1983), Pyszczynski and Greenberg (1987), Kunda (1990, 1999) and Baumeister and Newman (1994) have proposed models of the human inference process and how it may be susceptible to bias. Baumeister and Newman's (1994) approach will be described in some detail to serve as a framework for a discussion of empirical evidence bearing on biased information search, evaluation and integration processes. Baumeister and Newman's (1994) model outlines the various mechanisms (and the time at which they operate during the decision process) by which individuals regulate their own inference or decision process.

According to Baumeister and Newman, individuals engage in self-regulation when their decision process is motivated. Motivation, in their view, is not just a biasing factor but more generally a "source of efforts to guide the inference or decision process in any way - even to make it less biased and more accurate" (p. 3). The latter refers to the case where a lack of motivation or effort might result in a biased conclusion and the individual tries to avoid this happening.

The model proposes that making an inference or a decision involves four stages (see Table 4.1): (1) *gathering evidence* (evidence is assumed to exist as information in long-term memory or the immediate environment), (2) *assessing the implications of the evidence* (this is assumed to be an automatic process), (3) *re-assessment of the implications* (e.g., evaluating their validity, clarity, or strength, or the degree of their relevance; this may result in the rejection of some of the implications arrived at during the second stage), and (4) *integration of the results of the third stage* (e.g., by resolving inconsistencies or assigning relative weights to different factors). This sequence is not to be regarded as rigid; operations associated with a particular stage can begin before the operations of the previous stage have been completed and a completed stage can be re-initiated, for example, if the conclusion that was reached was unsatisfactory. The second stage, the assessment of implications of the available evidence, is assumed to automatic and, therefore, immune to self-regulation, but the subsequent stage during which implications may be re-assessed allows results of this automatic process to be overridden by a controlled process. Baumeister and Newman point out, however, that the automatic assessment of implications puts an important constraint on the self-regulation of inference and decision processes, since it prevents the individual from drawing any conclusion they like from whatever evidence available.

Baumeister and Newman distinguish between two basic motives underlying self-regulation, a need to reach an *optimal* conclusion (the correct or ideal one) and a need to reach a *particular* conclusion (see also Kruglanski, 1980, 1989, 1990; Kunda, 1990, 1999). Kunda (1990) calls the former *accuracy goal* and the latter *directional goal*. Empirical evidence suggests that these two motivational patterns underlie most inference and decision processes (e.g., Chaiken, Liberman, & Eagly, 1989; Pyszczynski & Greenberg, 1987, Larrick, 1993). As described in the previous chapter, they have also been found to explain accountability effects, with accountable decision makers wanting to be more accurate but also wanting to make a

decision that is easy to justify to their audience. According to Baumeister and Newman, the individual who wants to be accurate acts like an *'intuitive scientist'* (Fiske & Taylor, 1984), whereas the individual who wants to reach a particular conclusion acts like an '*intuitive lawyer*'. If one wants to arrive at a particular conclusion, for example, that one's cousin should get the position one has to fill, it is readily apparent why self-regulation of the decision process is necessary; it has to be guided towards the preferred conclusion. Like a lawyer who will try to build the best possible case for a client, irrespective of whether the client is innocent or not, the individual must try to assemble the best available evidence for his or her preferred conclusion or against an unwanted conclusion. Therefore, the desired outcome guides the self-regulation process. If the individual wants to maximise accuracy, on the other hand, self-regulation will focus on the objectivity of the decision *process* rather than its outcome, for example, by trying to make sure that all relevant evidence is considered in an unbiased manner. Baumeister and Newman state that "to the extent that the preferred conclusion is not the same as the accurate or optimal conclusion, self-regulation may serve ends of self-deception" (1994, p. 5). However, they also point out that self-deception fails, if it is acknowledged too openly. Therefore, self-regulation must be subtle or covert.

Not all decisions are assumed to be subject to self-regulation. Since self-regulation requires a fair amount of cognitive effort, only when decisions are complex and when the decision maker's personal involvement or the perceived importance of the decision is high enough to justify this effort, will the decision maker engage in self-regulation. A number of self-regulation mechanisms related to the different stages of the inference or decision process are available to the intuitive scientist and intuitive lawyer to achieve their respective processing goals. An overview is shown in Table 4.1.

When assembling evidence, the decision maker can control the amount, range, type, and valence of information that enters the decision process. Whereas the intuitive scientist will try to ensure that all relevant information will be searched and receive equal attention, the intuitive lawyer will search for information that favours his or her preferred conclusion and/or show selective inattention to contrary information, that is, will show a confirmation bias. Empirical evidence for such processes will be discussed later in this chapter. Attention can also be manipulated by changing the viewpoint one adopts and the time one spends on processing each piece of information. In addition, since the information search process often

does not have a natural endpoint, deciding when to stop searching for further information may be used to regulate the decision process (a so-called process of freezing, Kruglanski & Ajzen, 1983; Kruglanski, 1980, 1989, 1990). It is important to recognise, though, that a thorough information search does not guarantee that bias will be eliminated and that a person in the mindset of an intuitive lawyer does not necessarily search less information than someone in the mindset of an intuitive scientist.

Table 4.1. Mechanisms for self-regulation of inference (based on Baumeister & Newman, 1994, p. 6)

Decision Stage	Intuitive Scientist	Intuitive Lawyer
1. Gathering evidence	Thorough search, considering the opposite	Confirmation bias, selective inattention
2. Assessing implications	(Automatic)	(Automatic)
3. Re-assessment	Search for possible bias, adjustment, re-computation	Controlled override of unacceptable conclusions, selective criticism of unwelcome evidence
4. Integration	Rules and criteria set in advance	Rules and criteria set to emphasise favourable evidence

Assessing the implications of evidence that has been collected, as mentioned above, is an automatic process and therefore is not open to self-regulation, but the decision maker can exert some control over the results of this process by carrying out a re-assessment. Baumeister and Newman suggest that, at this stage, the intuitive scientist who wants to make an accurate decision will try to correct for any biases and distortions in the available evidence, provided they have been identified[1]. According to Strack (1992), a correction of bias may be achieved by either adding on an adjustment or by repeating the decision process with revised evidence. Making an adjustment is less effortful, but requires some knowledge about the size and the direction of the distortion. Since people often lack this knowledge, an over-adjustment (e.g., Hatvany & Strack, 1980) or under-adjustment (e.g., Ross, Lepper, & Hubbard, 1975) is likely.

[1] It has been shown repeatedly, however, that people are often not very good at recognising flaws in the evidence they have assembled (e.g., Fischhoff, 1982).

Baumeister and Newman suggest that the amount of processing that occurs between encountering a biased piece of evidence and becoming aware of the fact that it is biased is critical for how adequate an adjustment will be. The more information is processed, the more difficult it is to reverse the biasing effect. The intuitive lawyer, in contrast, uses the re-assessment stage for introducing a directional bias into his or her information processing, by, for example, selectively doubting the validity, clarity, strength or relevance of opposing evidence. He or she may also make inappropriate adjustments for information which is discredited or perceived to be biased. Again, the evidence for such confirmatory bias in information evaluation will be discussed in some detail later.

Baumeister and Newman argue that, while it may be difficult for the intuitive lawyer to search information in a confirmatory manner (since it is often not immediately obvious whether information is supporting or not before looking at it), the re-assessment of implications is easy to regulate, because the evidence is there and can be worked upon. In contrast, for the intuitive scientist, his or her accuracy goal is more difficult to achieve in the re-assessment stage than in the information search stage, because it is difficult to determine which flaws in the evidence, if any, are most problematical. As a result, the intuitive scientist is likely to concentrate on regulating information search, whereas the intuitive lawyer is likely to concentrate on the re-assessment of implications.

The final step of the decision process requires the combination and integration of evidence to arrive at an overall decision or conclusion. This may be difficult, because in many cases the assembled evidence may not be unequivocal but will support different decisions or conclusions (see the discussion of multiattribute decisions in Chapter 1). Procedures for resolving conflicting implications may involve weighting pieces of evidence differently and following certain decision rules. Baumeister and Newman state that an important difference between the intuitive scientist and the intuitive lawyer in this respect is the time at which such rules will normally be set. The intuitive scientist will set weights and the decision rule, including any cut-off points it requires (e.g., when the decision maker uses a conjunctive, disjunctive, or elimination-by-aspects rule), before collecting the evidence. The intuitive lawyer, on the other hand, will prefer to have evidence collected before applying rules that

guarantee the desired outcome[2]. Svenson's (1992, 1996) differentiation and consolidation theory, which will be discussed later in this chapter, spells out the different ways of supporting one's preferred conclusion in more detail.

Baumeister and Newman stress that the proposed sequence of information processing steps in their model is theoretical, in reality human inference is likely to be an on-line process; people will reach preliminary conclusions or make tentative decisions before having completed the information search and information evaluation stage. These preliminary decisions have also been referred to as hypotheses by other authors (e.g., Kruglanski & Ajzen, 1983; Pyszczynski and Greenberg, 1987). Intuitive scientists and intuitive lawyers differ in that the latter have a hypothesis before even starting to process information, whereas the former develop a hypothesis only during processing. The preliminary conclusion or hypothesis is updated with each new piece of information encountered. The updating process is another source of bias, since people typically do not make normative adjustments (e.g., Borcherding, 1981). Not only are the different processing operations performed almost simultaneously, but they must also be thought of as iterative, in that their result, that is, the decision or conclusion arrived at, may prompt the decision maker to return to previous steps. If the intuitive lawyer is unable to reach his or her preferred conclusion, he or she may search for more, preferably supporting information, and process it in such a way that the preferred conclusion is supported, or he or she may return to the re-assessment stage and change the assessment of unfavourable implications.

In conclusion, Baumeister and Newman's (1994) model of self-regulated inference provides an interesting theoretical framework for the analysis of the effects of different processing goals on information processing, in that it proposes that different processing goals affect different stages of the judgement and decision process. In particular, the model suggests that a decision maker with an accuracy goal who, according to Baumeister and Newman, will adopt the mindset of an intuitive scientist, will focus on regulating his or her information acquisition process by conducting a particularly thorough information search, while a decision

[2] This is unless the individual has prior expectations of how the evidence will turn out. In this case, rules can be set in advance and an impression of fairness and objectivity created. Baumeister and Newman give the example of voting laws which required people to pass a means or literacy test to be allowed to vote, in the expectation that unwelcome, disadvantaged groups of the population would be less likely to pass them.

maker with a directional goal, who will adopt the mindset of an intuitive lawyer, will focus on the evaluation and integration stage and try to achieve the desired conclusion by assessing information in a biased manner as well as employing decision rules that favour the desired outcome. The individual acting as an intuitive lawyer may also show a biased information search process if it is obvious to him or her which information supports the desired outcome. The processes of biased information search, evaluation and integration by which the individual attempts to reach a desired conclusion are of particular interest when trying to understand how accountable decision makers arrive at a decision that they consider acceptable to the people they are accountable to. We will therefore continue with a review of the literature on confirmation bias at these different stages of the judgement and decision process, starting with confirmation bias in information search.

Confirmation bias in information search

Confirmation bias, that is, the tendency to look for evidence that confirms a hypothesis or decision rather than to look for falsifying evidence, has been widely documented in different sub-disciplines of psychology (see Oswald, 1993, and Klayman & Ha, 1987, for a review). There are, however, at least two different ways in which the term has been understood by researchers in the past (Fischhoff & Beyth-Marom, 1983; Klayman & Ha, 1987; Skov & Sherman, 1985). Firstly, it has been used to refer to a tendency to search information that has a high likelihood of supporting one's hypothesis or preferred conclusion. This type of confirmation bias, which may be regarded as a hypothesis preservation or protection bias (Devine, Hirt, & Gehrke, 1990), has been demonstrated, for example, in studies conducted within the framework of dissonance theory (Festinger, 1959, 1964), where participants' post-decisional information search has been found to be biased towards information they expect to support their previously made decision (see Frey, 1986, for a review). Such effects will be discussed in more detail below. Secondly, the term 'confirmation bias' has been used to describe a tendency to test cases that are expected (or known) to be true under the working hypothesis, that is, to ask questions in which a 'yes' or positive answer would confirm one's hypothesis. Klayman and Ha (1987) have re-named this phenomenon the positive test strategy.

Hypothesis preservation bias

Festinger's (1957, 1964) theory of cognitive dissonance

Festinger's (1957, 1964) theory of cognitive dissonance has provided one of the major social psychological frameworks for systematically exploring biased information search in the form of a hypothesis preservation bias. Dissonance theory is based on the assumption that individuals strive for consistency in their cognitive system. Cognitions as the building blocks of the cognitive system can either be irrelevant to each other when there is no logical connection between them (e.g., "I smoke" and "Rome is a nice city"), or they can be relevant to each other (e.g., "I smoke" and "Smoking is bad for my health") when one has implications for the other. Only relevant cognitions are of interest for the theory. Relevant cognitions are said to be consonant when they are consistent with each other, for example, "I haven't got much money" and "I won't go on holiday this summer". A state of dissonance arises whenever two relevant cognitions are inconsistent with each other, that is, when considered by themselves, one of them follows from the opposite of the other. The above-mentioned cognitions "I smoke" and "Smoking is bad for my health" may serve as an example. Dissonance, according to Festinger's theory, is accompanied by negative arousal, which motivates the individual to reduce dissonance. This motivation increases with dissonance strength. Dissonance strength is defined as the ratio of the number of dissonant to the total number of relevant (i.e., consonant plus dissonant) cognitions, with each cognition weighted for its importance to the person. This implies that dissonance reduction can be achieved in a number of ways, such as increasing the number and/or importance of consonant cognitions or reducing the number and/or importance of dissonant cognitions. Generally, how exactly dissonance is reduced depends on the resistance to change of the various cognitions involved, with less resistant cognitions being more likely to be changed.

Festinger's theory generated a large amount of research, from which it soon became apparent that certain aspects of the theory had to be clarified and/or to be revised. The first revision was suggested by Brehm and Cohen (1962), who concluded that two conditions which Festinger had not mentioned were crucial for the occurrence of dissonance: freedom of choice and commitment to the decision. Only when the individual has been given a choice to perform a certain behaviour and committed him/herself to this behaviour, can dissonance

arise and dissonance reduction be observed. Cooper and Fazio (1984), in their new-look theory, pointed out that in order for dissonance to occur, apart from a situation of free choice, the individual also has to feel personally responsible for foreseeable, potentially aversive consequences.

Selective exposure research

Dissonance theory has been applied to a wide range of phenomena, first and foremost to attitudinal change after attitude-inconsistent behaviour (forced compliance). The applications that are particularly relevant to the work presented here, however, are post-decisional preference changes and selective exposure to information. The latter involves the search of supportive information after a decision or a tentative decision has been made. According to dissonance theory, whenever a decision is made, dissonance occurs because the decision maker anticipates being stuck with the negative aspects of the chosen alternative while losing the positive aspects of the non-chosen alternative. The 'selective exposure' hypothesis (Festinger, 1957) predicts that the individual will try to reduce this dissonance by looking for information that is consistent with his or her decision and avoid information that is inconsistent with this decision. Experimental studies on mainly post- and occasionally pre-decisional information search conducted within the framework of dissonance theory typically present participants with a number of article or essay titles to choose from. The titles strongly suggest whether the information provided is supportive or non-supportive of a certain decision. Participants indicate the articles (usually a fixed number) they would like to read and/or they would not like to read. While this allows testing the selective exposure hypothesis, it does not allow a detailed analysis of the decision process.

Although there has been some empirical support for the selective exposure hypothesis, overall the evidence for it is mixed at best (see Frey, 1986, for a detailed review). Over the years it has emerged that the original selective exposure hypothesis was far too simplistic and that a number of different variables moderate the expected effect. Early studies often not only failed to find a preference for consonant information, but even demonstrated a preference for dissonant information (e.g., Freedman, 1965a; Sears, 1965). These and the results of other experiments led Freedman and Sears (1965) to conclude that dissonance theory did not apply to post-decisional information search.

The initial lack of empirical support for the selective exposure hypothesis induced Festinger (1964) to revise it and specify three conditions under which people adopt a so-called approach-avoid strategy, that is, attempt to reduce dissonance by counter-arguing dissonant information rather than looking for consonant information, exposing themselves to dissonant information in the process: (1) when dissonant information is perceived as easily refutable, (2) when avoiding dissonant information will cause even greater dissonance in the future, and (3) under high dissonance, when a revision of the decision is possible.

The *ease of refutability* of information may be argued to depend on three factors (Frey, 1986), the strength of the argument, the competence or credibility of the source, and the stability of the decision maker's cognitive system. Accordingly, a preference for dissonant information can be observed when dissonant arguments are weak (Lowin, 1967; Kleinhesselink & Edwards, 1975), when they are attributed to a source of low competence and/or low credibility (Lowin, 1969; Frey, 1981a), and when the decision maker has previously been exposed to consonant information (Schwarz, Frey, & Kumpf, 1980; Frey, 1981a).

The second central factor in Festinger's (1964) re-formulated theory, the *long-term avoidance of dissonance*, also received empirical support. Dissonant information is preferred over consonant information if it is perceived as useful for avoiding dissonance in the long run. This is, for example, the case if similar decisions are anticipated or when a revision of the original decision is imminent (Frey, 1981b). Dissonant information can also be perceived as useful if participants anticipate having to justify their decision in a debate, as Canon (reported in Festinger, 1964) demonstrated. In his experiment, participants were either told that they would have to defend their decision in a debate or that they would have to write an essay defending their decision. After the decision and before the justification, they could choose to read up to three articles out of a set of five (2 consonant, 2 dissonant, 1 neutral). Canon predicted that participants in the 'debate' condition would prefer dissonant information in order to give a favourable self-presentation, whereas participants in the 'essay' condition would prefer consonant information in order to present a coherent argument. These predictions were confirmed. Another study that investigated the effect of utility of information as a consequence of having to justify a decision was one by Freedman (1965b). He employed three different utility conditions: a condition where consonant information was

assumed to be more useful than dissonant information (in which participants had to justify their decision in an essay), a condition where dissonant information was assumed to be more useful than consonant information (in which participants expected to be presented with opposing arguments which they had to refute in writing), and a condition where consonant and dissonant information were assumed to be equally useful (in which participants had to evaluate decisions made by others). Consistent with predictions, it was found that the type of information preferred by participants reflected the utility of information. These studies obviously have important implications for any predictions of the effects of accountability on information search and will be returned to later.

Festinger's (1964) final point of revision of his original theory relates to *dissonance strength* affecting selective exposure. The tendency to look for consonant information should increase with dissonance strength up to the point when dissonance becomes so strong that the individual contemplates a change of his or her decision and therefore starts searching for information that supports a different decision. This implies a curvilinear relationship between dissonance strength and selectivity, with dissonance reduction peaking at a moderate level of dissonance. An important qualification, however, is that the decision must be reversible. If the decision is irreversible, a linear increase of consonant information search can be expected (Frey, 1981b). In a study by Frey (1982), participants had to choose the role of player A or B in a 30-trial card game for money. They were told that they could change their decision after the twelfth trial, but had to pay a small amount of money for it. After twelve trials, participants were asked to select from twelve short essays allegedly written by previous players (6 of which supported the A choice and 6 of which supported the B choice) the three essays they would most like to read and the three essays they would not like to read. As expected, preference for consonant information and avoidance of dissonant information was largest when participants had neither won nor lost a large amount of money. As gains and losses increased the preference for consonant information decreased. Also, participants' ratings of the probability of changing their position for the next 18 trials revealed that the higher the probability of revision, the less consonant information was chosen and the less dissonant information was avoided. Frey proposes that, in this situation, two processes operate in parallel: a tendency to seek consonant information, in order to reduce dissonance, and a tendency to seek dissonant information, in order to make the right decision about

whether to change role. As losses increase, the latter is assumed to override the former, leading to a net decrease in preference for consonant information. Note that Frey's assumption points to the basic motivational conflict underlying much of human information processing, the conflict between a need to confirm one's actions and a need for accuracy (see Baumeister & Newman's model above).

Another situation where dissonant information is preferred to consonant information is when a *norm of fairness and/or impartiality* is salient (Canon, 1964; Freedman, 1965a; Frey, 1991). This may be a particularly appropriate explanation for Freedman's (1965a) and Sears's (1965) findings which failed to support the selective exposure hypothesis and which were obtained in a job selection and jury context, respectively. It may be argued that in such situations, possibly because of the implied severity of the consequences of their decision, participants have a particularly strong need for accuracy, hence, they search not only consonant but also dissonant information. Finally, Canon (1964) named *curiosity* as a factor that can obscure selective exposure effects.[3]

In contrast, a selective exposure effect is likely to be obtained, when there are implications for participants' self-esteem. Frey (1981c) and Frey and Stahlberg (1986), for example, found that participants who had been given false negative feedback about their performance in an intelligence test, showed a preference for information that argued against the validity of intelligence tests over information that argued for their validity. In Frey and Stahlberg's (1986) second study, this effect was largely diminished, however, when there was an opportunity for an external attribution for failure, when participants had been told that their test may have been scored incorrectly. Holton and Pyszczynski (1986) showed that

[3] There are a large number of other variables which have been found to affect selective exposure. These will not be discussed in detail, however, as such a discussion would go beyond the scope of this work. They include variables such as *participants' confidence in their decision* (low confidence results in stronger selective exposure, Canon, 1964), *public commitment* (a public commitment to a certain opinion results in stronger selective exposure, Schwarz, Frey, and Kumpf, 1980; Sweeney & Gruber, 1984, Frey & Stahlberg, 1986), the *amount of each type of information* from which to choose (if there is more than one item of dissonant information, participants prefer consonant over dissonant information, Frey, 1985), whether information incurs a *cost* or not (if it does, selective exposure is stronger, Frey, 1981c), perception of dissonant information as *valid* or *reliable* (this increases the tendency to expose oneself to consonant information, Frey and Stahlberg, 1986), and, finally, *personality variables*, for example, dogmatism (high dogmatism results in stronger selective exposure, Clarke & James, 1967) and repression (repressors search more consonant information than sensitizers, Olson & Zanna, 1979).

participants who had been evaluated favourably by a confederate searched for more information when they expected it to suggest that the confederate was a competent evaluator, whereas participants who had received an unfavourable evaluation searched for more information when they expected it to suggest that the confederate was an incompetent evaluator. Participants, however, who had merely observed the confederate's evaluation of another participant, did not show such bias. Indeed, Steele (1988; Steele & Liu, 1981, 1983) has argued that individuals engage in dissonance reduction to re-affirm that they have behaved in accordance with positive values. Note that the motive underlying dissonance here is assumed to be self-integrity (i.e., a desire to maintain a *positive* view of the self) rather than self-consistency.

The discussion so far has only covered post-decisional information search. Pre-decisional information search has rarely been investigated within the framework of dissonance theory. This may be because Festinger himself stated that "The person .. continues to seek new information ... This continued information seeking and information evaluation remains, however, objective and impartial" (1964, p. 152). An experiment by Jecker (reported in Festinger, 1964) indeed seemed to suggest that pre-decisional information search was unbiased, since the time participants spent reading supporting information before they chose a partner for a game did not differ significantly from the time they spent reading non-supporting information. Participants who were given the opportunity to read information only after they had already made their choice, on the other hand, spend more time reading supporting information than they did reading non-supporting information. It may be argued, however, that time spent reading information is not a good indicator of dissonance reduction. A longer time spent reading information is usually interpreted as a stronger desire on the part of the individual to expose him- or herself to this information. An argument could also be made for the exact opposite, however. As Festinger (1964) himself has suggested, if the individual knows that there is information which is dissonant with his or her decision, the best way of reducing dissonance is to re-interpret this information. In order to be able to do so, however, the individual must expose him- or herself to the dissonant information and may be expected to spend a longer time on the cognitively effortful process of re-interpretation than in a case of simple information avoidance (Grabitz & Grabitz-Gniech, 1973). Another major difficulty one faces when analysing pre-decisional information search, is to identify the point at which

the decision maker commits him- or herself to an alternative, and thereby, it may be argued, switches from pre- to post-decisional information search. Some authors (e.g., Janis, 1959) have dealt with this problem by defining a decision as an explicit overt action which signals a commitment to carrying out a specific task or adhering to a particular course of action. This allows a clear distinction between the pre- and post-decisional phase. As Grabitz and Grabitz-Gniech (1973) point out, however, it is questionable whether the point at which a decision is made public accurately reflects the timing of significant psychological change in the accompanying cognitive process. Rather, it is likely that any explicit decision is preceded by an implicit, tentative one. Given that it lies in the very nature of tentative decisions that they can be changed, the goal of information search may change at any time and may mask confirmation bias.

From the discussion above it should have become evident that the processes governing pre- and post-decisional information search are far from being straightforward. A large number of moderating variables have to be taken into account when trying to predict whether consonant or dissonant information will be preferred. Wicklund and Brehm (1976, see also Frey & Wicklund, 1978; Frey, 1986) suggested that the reason for this may be the fact that selective exposure requires new behaviour, namely actively attending to previously unexplored information, evaluating it and perhaps even learning it. This is more complex than, for example, a change in one's evaluation of information, which only requires changing something which is already part of one's cognitive system and where dissonance reduction can be generated from within. Given that selective exposure requires additional behaviour, it is likely that factors other than the existence of dissonance have an impact on it as well.

Whereas the tendency to search for consonant information has received empirical support under certain specified conditions, Frey (1986) concludes that the evidence for the other part, relating to the avoidance of dissonant information, is even weaker. Among the exceptions is a study by Baumeister and Cairns (1992), who showed that participants coped with an unfavourable personality evaluation by reducing the amount of time they spent reading it and subsequently showed impaired memory for it. Also, Sweeney and Gruber (1984) showed that participants who supported President Nixon tended to avoid news on the Watergate affair, whereas participants who opposed Nixon followed the Watergate news closely. Information avoidance, if it is observed at all, often does not vary as a function of the

previously discussed variables, however. Frey suggests that this may be, because the avoidance of dissonant information does not reduce dissonance, it just does not increase it. Also, a general avoidance of dissonant information may not be effective in terms of long-term avoidance of dissonance. A methodological explanation for the failure to observe dissonance avoidance has been proposed by Götz-Marchand, Götz, and Irle (1974). Usually, selective information seeking is measured before selective information avoidance. It may be that by the time selective avoidance is measured, dissonance has already been reduced sufficiently and that there is therefore no need for the individual to attempt further reduction by selectively avoiding dissonant information. This assumption was supported by Götz-Marchand et al.'s finding that the opportunity to reduce dissonance by negatively evaluating an intelligence test in which participants had allegedly performed badly, was used more often when it was the first dependent measure compared to when it appeared in fifth position in the post-test.

Other hypothesis preservation paradigms and their link to accountability research

The discussion of evidence in support of confirmatory information search so far has heavily relied on research within the framework of dissonance theory, but this is by no means the only research programme that has provided such evidence. A hypothesis preservation bias can not only be observed for externally available information but also for information that is searched in memory. Kunda and her co-workers (see Kunda, 1990, 1999, for a review) have repeatedly demonstrated a biased memory search. For example, participants motivated to believe that they possess desirable intellectual or personality traits search for and report autobiographical memories that are consistent with those traits (Kunda, 1987; Sanitioso, Kunda, & Fong, 1990). Among the many other approaches to hypothesis preservation bias, the comprehensive reporting of which is beyond the focus of this work, Kruglanski's (1980, Kruglanski & Ajzen, 1983, Kruglanski & Klar, 1987) theory of lay epistemology deserves special mention, as it is consistent with Baumeister and Newman's model of self-regulated decision processes and can be related to information search behaviour observed in selective exposure studies. Kruglanski postulates three personal motives that affect information processing. Firstly, a fear of invalidity, which means that the individual is afraid of making incorrect judgements and therefore hesitates to stop information search. This is similar to the

mindset of the intuitive scientist with an enhanced need of accuracy who searches a lot of information, and may explain the search of dissonant information after consonant information has been exhausted. Secondly, a need for structure, which results in a tendency to stop information search, in order to establish a structured and predictable world, and thirdly, a need for specific conclusions, which results in a discontinuation of information processing as soon as the desired conclusion is reached. The latter two factors both closely resemble the motives of the intuitive lawyer and may explain a preference for consonant information that increases cognitive consistency and supports the individual's preferences. According to Kruglanski, information search should not be biased when fear of invalidity is high (i.e., accuracy goals are strong), whereas the tendency to look for supporting more than non-supporting information should be observed with a high need for structure or a strong preference for specific conclusions (i.e., strong directional goals). This has indeed been observed (e.g., Kruglanski & Ajzen, 1983; Kruglanski & Freund, 1983). In studies carried out to test Kruglanski's model, a typical way of inducing high fear of invalidity has been to tell participants that they have to justify their judgements to others or that their judgements will be publicly compared to the 'correct' answer. This is virtually identical with common accountability manipulations and supports the suggestion that accountability results in strong accuracy goals that deepen information search. Although a need for structure and a preference for specific conclusions have not been explicitly investigated within the context of accountability, it may be argued that making people accountable to an external audience with known views or making them accountable after they have already committed themselves to a particular option, will heighten their need for a confirmatory information search that supports the conclusion expected to be preferred by the external audience or arrived at earlier.

Positive test strategy

Demonstrations of the positive test strategy as the second type of confirmation bias have been, for example, provided by Wason (1960) and others employing variations of his rule discovery paradigm. In the 2-4-6-paradigm, participants are presented with a triple of numbers (2, 4, and 6) which is said to comply with a rule that participants are meant to discover. In order to discover the rule, participants have to generate new triples on which they receive feedback from the experimenter as to whether these triples conform or do not conform

to the rule. Participants typically quickly form an initial hypothesis, such as 'three consecutive even numbers', and generate triples which fit this hypothesis, such as 6, 8, 10, but fail to test triples that would falsify their hypothesis. If the hypothesised rule is embedded within the correct rule (e.g., 'any sequence of three increasing numbers'), participants employing the positive test strategy receive continuous positive feedback which increases confidence in their hypothesis but prevents them from discovering the correct rule. This effect is very robust and has been replicated many times (e.g., Beyth-Marom & Fischhoff, 1983; Doherty, Mynatt, Tweney, & Schiavo, 1979; Klayman & Ha, 1985). Another example of a positive test strategy may be seen in people's tendency to weight instances in which a cue is present more heavily than instances in which a cue is absent when making judgements about the contingency between a cue and an outcome (Alloy & Tabachnik, 1984; Arkes & Harkness, 1983; Lipe, 1982; Nisbett & Ross, 1980; Schustack & Sternberg, 1981). A positive test strategy differs from hypothesis preservation bias in that, although participants search for information which, if found, confirms their hypothesis, they have no underlying motivation to preserve or protect this hypothesis.

A clear distinction between the two types of confirmation bias is not always possible, however (Oswald, 1993; Skov & Sherman, 1986). Snyder and co-workers (Snyder & Swann, 1978a, b; Snyder & Uranowitz , 1978; Snyder & Cantor, 1979; Snyder & Campbell, 1980; see Snyder, 1981 for a review) investigated how individuals tested trait hypotheses about people with whom they anticipated social interaction. Participants were, for example, told that they had to judge the extent to which a target person's behaviour was either introvert or extravert. In order to do so, they were asked to select 12 questions from a pool of 26, these were either directed at introversion (e.g., "What do you dislike about loud parties?"), extraversion (e.g., "What would you do if you wanted to liven things up at a party?") or were neutral ("What do you think the good and bad points of acting open and friendly are?"). Snyder and colleagues found evidence for a confirmation bias, in that participants mostly selected questions which matched the hypothesis they were testing, that is, preferred questions that assumed the person was introvert when they had to judge the extent of introversion, and questions that assumed the person was extravert when they had to judge the extent of extraversion (see also Evans, 1989). This may be interpreted as evidence for a positive test strategy, but could also be regarded as a hypothesis preservation bias, in that the selected questions had a higher á priori

likelihood of confirming participants' hypothesis than those not selected. The nature of the questions participants had to choose from, however, does not allow to distinguish between a positive test strategy and a hypothesis preservation bias, as the questions were always low in diagnosticity (the probability of getting a positive answer from someone possessing the trait in question was very similar to the probability of getting a positive answer from someone not possessing the trait), and any answer could have been taken as supporting the hypothesis. For example, no matter whether it is introverts or extraverts who answer the question about what they dislike about loud parties, they will always list dislikes and therefore appear introvert (Skov & Sherman, 1986). A stringent test between a positive test strategy and a hypothesis preservation bias requires manipulation of diagnosticity, since a true hypothesis preservation bias can only be said to be present if participants prefer questions that are low in diagnostic value but high in likelihood of confirming their hypothesis over questions that are high in diagnostic value but low in likelihood of confirming their hypothesis. This point was made by Trope & Bassok (1982, 1983; Bassok & Trope, 1984) who, using a similar paradigm as Snyder and Swann, found that participants showed a preference for questions that were high in diagnostic value regardless of whether they were hypothesis-confirming or not. This finding seems to imply that Snyder and Swann's results were based on a methodological artefact and that confirmatory information search is not nearly as pervasive as they claimed. Both Skov and Sherman (1986) and Devine, Hirt, and Gehrke (1990), however, were able to show that among questions with equally high diagnostic value, participants prefer those that are hypothesis-confirming. Also, Hodgins and Zuckerman (1991) found that when participants had to generate questions themselves rather than select them from a list of pre-formulated ones, they tended to ask about features that were consistent with their hypothesis. Participants' inferences drawn from hypothetical answers to their self-generated questions also showed a greater tendency to confirm the hypothesis they were testing than the alternative hypothesis and a greater tendency to confirm the hypothesis following hypothesis-consistent compared to hypothesis-inconsistent questions. This tendency for hypothesis preservation seems to be very robust and does not disappear when people are made accountable. In a study by Strohmer and Shivy (1994), in which counsellors were asked to select the ten most important facts from a client narrative, in order to test whether the client lacked self-control, and were either told or not told that they would have to explain their information selection to

another counsellor, they selected more confirmatory than disconfirmatory information, regardless of whether they had been made accountable or not.

It should be pointed out here that a confirmation bias in form of a positive test strategy can be functional and does not necessarily carry negative consequences (Klayman & Ha, 1987). Important for whether or not a positive test strategy is functional is (1) whether the context in which a decision is made carries more concern for false positives or for false negatives, and (2) which strategy will result in more conclusive falsifications. In a personnel selection context, for example, it is more important to avoid selecting someone who is not suitable for the job (i.e., to avoid false positives) than to avoid missing out on someone who would have been suitable (i.e., to avoid false negatives). In such a situation, a positive test strategy (testing applicants who are assumed to be suitable) is beneficial, because it helps to identify false positives. In contrast, in a situation where one has to identify people with an infectious disease, for example, it would presumably be more important to avoid undiagnosed cases which can contribute to the spread of the disease than to avoid treating someone unnecessarily. In this case, a negative test strategy would be better. Also, if the number of potentially positive instances is small, a positive test strategy may help to keep the number of tests down (e.g., one only has to search information for the small number of potentially successful job applicants instead of the large number of potentially unsuccessful ones) and therefore minimises the cognitive effort individuals have to expend. Finally, it can be shown that in such a situation, the probability of receiving conclusive falsifications is higher for a positive than for a negative test strategy and a positive test strategy would therefore be more effective (Klayman & Ha, 1987).

Summary

The review of the literature investigating confirmation bias in information search has suggested that there are two different manifestations of such bias, a tendency to search information that has a high likelihood of supporting one's hypothesis or preferred conclusion and a tendency to test cases that are expected (or known) to be true under the working hypothesis. The former has been named hypothesis preservation bias, the latter positive test strategy. Whereas there seems to be strong evidence supporting the widespread use of a positive test strategy, evidence for a hypothesis preservation bias, especially when it has been

collected in studies employing a selective exposure paradigm, is much more inconclusive and suggests that there are also various factors that encourage a less biased information search, including the ease of refutability of non-confirming information, the possibility of long-term avoidance of dissonance, the possibility of eliminating high dissonance by reversing one's decision, and a norm of fairness and/or impartiality. All these factors are consistent with the idea that strong accuracy goals (or a high fear of invalidity, in Kruglanski's terms), as they are typically encountered in individuals with the mindset of an intuitive scientist, encourage a more exhaustive and therefore less biased information search. On the other hand, hypothesis preservation bias is likely to be observed when information search has strong emotional consequences (Pyszscynski & Greenberg, 1987), for example, because the individual's self-esteem is threatened and the desire to arrive at a particular conclusion is stronger than the one to arrive at an accurate conclusion. Both the tendency for non-confirmatory and confirmatory information search may be linked to accountability. A non-confirmatory information search may be expected if the individual is made accountable to an audience with unknown views, a confirmatory information search, on the other hand, if the individual is made accountable to an audience with known views or after he or she has already expressed a preference for a particular option.

Confirmation bias in information evaluation and integration

Confirmation biases may not only be observed in information search but also in information evaluation and integration, for example, when individuals reinterpret information in such a way that it supports their preferred conclusion and selectively doubt the validity of negative evidence (stage 3 of Baumeister & Newman's model) or place disproportionately high weight on information that supports the preferred conclusion (stage 4 of Baumeister & Newman's model). Again, dissonance theory (Festinger, 1957, 1964) has provided one but not the only theoretical framework that explains such biases. Dissonance theory assumes that, after they have made a decision, individuals attempt to reduce the dissonance that arises from the disadvantages of the chosen alternative and the advantages of the non-chosen alternative by increasing their evaluation of the chosen alternative and/or decreasing their evaluation of the non-chosen alternative(s). Brehm (1956) was one of the first researchers to demonstrate this effect. In his study, he asked female college students, who thought that they took part in a

market research study, to indicate how much they liked each of eight electrical household appliances. Later, they were told that, as a reward for their participation, they could choose one of two products to take home. For some participants, the products they could choose between differed in rated attractiveness by ½ to 1½ points ('high dissonance' condition), for others they differed by 3 points ('low dissonance' condition). A control group was given one of the products as a gift without any choice. After having made a choice, participants were asked to re-evaluate the products. As expected, participants in the experimental conditions increased their liking for the chosen product and decreased their liking for the non-chosen product (spreading apart effect), whereas in the control condition no significant change of evaluation could be observed. Also, the effect was stronger in the 'high dissonance' than in the 'low dissonance' condition. While this study is one of many that demonstrate post-decisional confirmatory information evaluation, this effect can also be observed before a decision is made (Boiney, Kennedy, & Nye, 1997; Russo, Medvec, & Maloy, 1996).

A large number of studies have investigated individuals' cognitive reactions to information they are motivated to evaluate negatively. In Frey's (1978), Wyer and Frey's (1983) and Pyszczynski, Greenberg and Holt's (1985) studies, for example, participants who had been given false positive feedback on their performance in a test rated a study that concluded that the test was highly valid as more convincing and better conducted than a study that suggested low validity of the test, whereas the opposite was true for participants who were led to believe that they had done poorly in the test. Similarly, Lord, Ross, and Lepper (1979) found that participants who either supported or opposed capital punishment rated a study that confirmed their belief as more convincing and better conducted than another that did not, irrespective of whether the study was cross-sectional or longitudinal. In Sherman and Kunda's (1989) investigation, participants motivated to disbelieve a study about the relationship between high caffeine intake and fibrocystic disease were less persuaded by it. This effect was shown to be mediated by a biased evaluation of the methods employed in the study. Participants who felt threatened because they drank at least two cups of coffee per day listed fewer strengths of the employed methodology than non-threatened participants, and rated various methodological aspects as less sound (see also Kunda, 1987; Liberman and Chaiken, 1992).

In line with the aforementioned results, there is also evidence that preference-inconsistent information leads to more cognitive activity than preference-consistent information. Gilovich (1983), for example, demonstrated that gamblers spend more time thinking about their losses than about their wins, because of trying to explain away their losses. Schaller (1992) found that participants, who had been randomly assigned to one of two groups which took an intelligence test, readily inferred that their own group was more intelligent when it performed better than the other group. When the other group did better, however, participants took the difficulty of the test items into account to arrive at a judgement. Similarly, Ditto, Scepansky, Munro, Apanovitch, and Lockhart (1998) demonstrated in a series of studies that participants' inferences drawn from favourable interpersonal feedback showed a correspondence bias, that is, failed to take situational constraints into account, whereas inferences drawn from unfavourable interpersonal feedback were sensitive to situational constraints. This sensitivity, however, disappeared under cognitive load. Finally, the tendency to apply different levels of stringency when evaluating preference-consistent and preference-inconsistent information can also be observed in people's use of stereotypes. Devine (1989), for example, suggested that both prejudiced and non-prejudiced individuals have racial stereotypes and that these stereotypes are activated with equal ease and speed. Unlike non-prejudiced individuals, however, prejudiced individuals do not engage in a re-assessment of the implications suggested by the activation of racial stereotypes and therefore express stereotypical views.

The fact that the validity of preference-consistent information is rather uncritically accepted, whereas preference-inconsistent information elicits cognitive activity in order to explain it away also implies that less preference-consistent than preference-inconsistent information is required to reach a conclusion. This 'quantity of processing' approach was supported in a series of studies by Ditto and Lopez (1992), who found that participants who were presented with information suggesting that a dislikeable student was less intelligent than a likable student requested less information to conclude that he was less intelligent than participants presented with information suggesting that the dislikeable student was more intelligent needed to conclude that he was more intelligent. Participants presented with preference-inconsistent information did not require more information to make their decision than participants without a particular preference, suggesting that it was the ready acceptance

of preference-consistent information (also referred to as self-enhancement bias, Miller & Ross, 1975) rather than an overly critical processing of preference-inconsistent information which caused the effect in this instance. A reluctance to accept preference-inconsistent information (self-protection bias) was observed in two further studies, where participants were found to take longer to decide that an unfavourable medical test result was complete, re-tested themselves more often, were more likely to state factors that could have affected test accuracy and rated test accuracy as lower than did participants who had obtained a favourable result. Ditto and Lopez's (1992) findings illustrate well how the processes of biased information search and biased information evaluation and integration interact with and mutually reinforce each other.

It should be mentioned again at this point that many of the above results may be explained in non-motivational terms. For example, the finding that favourable intelligence test performance feedback is accepted uncritically, whereas unfavourable feedback is attributed to a lack of validity of the test, may be explained by assuming that students, who typically serve as participants in such studies, usually have a history of successful academic performance. Unfavourable feedback therefore violates their expectations and needs an explanation, whereas favourable feedback confirms their expectations and does not call for an explanation. Hence, it is not necessary to assume that participants are motivated to enhance their self-image; an outsider who has no vested interest in the issue would be expected to arrive at the same conclusion. Indeed, many studies have demonstrated that biased evaluations can not only be observed when individuals have a self-related motivation to reach a particular conclusion but also when they just have certain prior expectations not related to personal goals. In a study by Darley and Gross (1983), for example, participants watched a videotape of a child whose academic ability they were asked to evaluate. The first section of the tape implicitly suggested that the child either came from a high or a low socio-economic background. Participants then either saw or did not see an inconclusive second section that showed the child taking an academic test. A fifth group just watched the second section of the tape. Participants who did not see the second part of the tape rated the child's ability level as about average for her age, regardless of whether they believed the child to come from a high or low socio-economic background. However, those participants who had seen the ability test and believed that the child came from a high socio-economic background rated the child's

ability well above average, whereas those who had seen the identical clip but believed her to come from a low socio-economic background rated her ability below average. Both groups referred to the ability test in order to support their respective views; they rated the difficulty of the test, the number of problems the child had answered correctly and the number of test behaviours that reflected either positively or negatively on the child's achievement level in such a way that their expectancy was confirmed. This suggests, consistent with cold cognition accounts, that a hypothesis confirmation bias also occurs when the individual has no personal motivation to arrive at a certain conclusion.

The biases discussed so far have largely focused on the re-evaluation of information. As suggested by Baumeister and Newman (1994), bias may also be introduced through the way in which information is integrated. One such mechanism described in the literature is the differential weighting of preference-consistent and preference-inconsistent information. In Boiney, Kennedy and Nye's (1997) study, for example, participants who were motivated to approve a new product implicitly weighted an expert forecast of sales that was considerably higher than those of three other experts more strongly than participants who were not particularly motivated to favour the introduction of the product, and, as a consequence, were more likely to advocate the introduction of the product. Dunning, Meyerowitz, and Holzberg (1989) found that when asked to rate themselves compared to their fellow students on various ambiguous traits, in order to achieve a desirable self-image, participants placed the greatest weight on a dimension on which they looked best. Similarly, in Dunning, Perie, and Story's (1991) study, when asked which traits were more important for a good leader, people-oriented participants rated interpersonal skills higher than goal-oriented participants, whereas goal-oriented participants rated ambition and persistence higher than people-oriented participants. Ginossar and Trope (1987) found that participants assigned to play the role of a lawyer failed to show the usual base rate fallacy (Kahneman & Tversky, 1973) and took base rate information into account if it supported the conclusion they were arguing for. Finally, studies conducted within the framework of Svenson's (1992, 1996) differentiation and consolidation theory (presented below) have demonstrated confirmative information integration in multiattribute decision making.

It is important to stress again at this point that individuals are not free to reach any conclusion they wish but are subject to reasonableness constraints. A compromise has to be

reached between the wish to reach a particular conclusion and the plausibility of this conclusion given the available evidence (Kunda, 1990, 1999, Boiney, Kennedy, & Nye, 1997). Given that individuals strive to uphold an illusion of objectivity (Pyszczynski & Greenberg, 1987), that is, pretend that they arrived at an accurate and objective conclusion, self-deception will only work if this illusion can indeed be upheld (Baumeister & Newman, 1994). Many studies have provided empirical support for this assumption. For example, in the study by Dunning, Meyerowitz, and Holzberg (1989) reported earlier, the tendency for participants to interpret ambiguous traits in such a way that their self-image was maximised was reduced when they were given particular evaluation criteria beforehand. Also, in Boiney, Kennedy, & Nye's (1997) study, participants only weighted the supportive forecast more strongly when this was necessary to arrive at the desired conclusion. A classical study by Snyder, Kleck, Strenta, & Mentzer (1979) also demonstrates individuals' awareness of justification constraints. Participants were asked to decide which of two rooms to wait in for their participation in an alleged experiment. In one of the rooms, a supposedly handicapped person (a confederate of the experimenter) was already waiting; the other room was occupied by a non-handicapped person. Also, in both waiting rooms either the same or two different TV programs were being shown. Snyder et al. (1979) found that when different TV programs were being shown in the two rooms, participants tended to choose the room that enabled them to avoid the handicapped person and justified their choice in terms of program preference. However, when the same TV program was being shown in both rooms and it was therefore not possible to use this fact as a justification for discriminatory behaviour, participants chose to wait with the handicapped person more often. Hsee (1995, 1996) has formalised the idea of justification constraints on confirmatory biases in multiattribute decision making in his theory of elastic justification. His concept of elasticity refers to uncertainty or ambiguity in attribute outcomes, for example, when only a range of possible outcomes instead of fixed outcomes is known. Given the choice between two options, one of which is superior on an attribute directly related to the task (justifiable factor) and the other of which is superior on a factor that is not central to the task but tempting to the decision maker (unjustifiable factor), the decision maker is assumed to first look to the justifiable factor to see whether there is a justification for making the desired decision and if that is not the case, take the unjustifiable factor into account only when there is elasticity in the justifiable factors. Hsee (1995) gives the example of a businessman who has to choose between two cities to go to for a business

trip to find as many buyers as possible for a product. In the first city, he expects to find more buyers than in the second city (justifiable factor), but the second city is more enjoyable (unjustifiable factor). According to the theory of elastic justification, the businessman will only choose the second city if the exact number of buyers he expects to find in either city is not exactly known and may be constructed to be higher for the second than for the first city. These predictions were confirmed in two series of experiments using different choice contexts (Hsee, 1995, 1996). Participants were found to be more likely to choose the alternative that was superior on the unjustifiable factor when there was uncertainty in the justifiable factor than when there was no such uncertainty, no matter whether the uncertainty was found in just one or in both alternatives. Another theory that specifies how individuals may use uncertainty in attribute outcomes to bolster their decision is Svenson's (1992, 1996) differentiation and consolidation theory, which will be outlined next.

Svenson's (1992, 1996) differentiation and consolidation theory

A theoretical framework that throws more light upon processes of confirmatory information evaluation and integration in a multiattribute decision context has been provided by Svenson (1992, 1996). His differentiation and consolidation (Diff Con) theory takes up and extends ideas from Festinger's (1957) dissonance theory and applies them to multiattribute decision situations. The theory states that during the process of making decision, alternatives are differentiated from each other. This is done in order to let one alternative appear sufficiently superior so that the decision for it can withstand future challenges, like unwanted negative consequences or changes in the decision maker's preference system. After the decision, consolidation processes follow to maintain or increase the achieved degree of differentiation. Differentiation and consolidation involve the same basic processes, the main difference between them is the time at which they operate; differentiation happens before and consolidation after a decision has been made.

In the pre-decision phase, the decision maker is assumed to eliminate alternatives until only two or three remain for further consideration and a tentative choice will be made. Differentiation processes are assumed to focus on the tentatively chosen alternative and its closest competitor. The aim is to increase the perceived difference in overall attractiveness to such an extent that the decision maker feels confident enough to make a final decision. This

aim can be achieved by either increasing the perceived attractiveness of the tentatively chosen alternative or by decreasing the perceived attractiveness of the competing alternative. Diff Con theory does not explicitly state the conditions under which one or the other is more likely to occur, however. Once a decision has been made, consolidation processes set in, in order to defend one's decision against anything that reflects unfavourably upon the chosen alternative and may lead to regret. Consolidation is assumed to happen regardless of how great the evaluative advantage of the chosen alternative already is. The theory thus makes an interesting prediction that deviates from predictions derived by similar theories, such as dominance structuring theory (Montgomery, 1983, 1989); differentiation and consolidation processes will also be observed when one alternative dominates the others, that is, its attribute outcomes are as good as or superior to those of the other alternatives.

Svenson (1992) lists a number of differentiation[4] processes, of which the so-called structural differentiation processes have received the most empirical attention, because they allow the clearest predictions. Structural differentiation is differentiation by means of changing the mental representation of decision alternatives and can be achieved by four different processes: (1) attractiveness restructuring, (2) importance differentiation, (3) fact restructuring, and (4) problem restructuring. *Attractiveness restructuring* involves changes in how attribute outcomes are evaluated. A negative outcome on an important attribute may be evaluated less negatively than before and/or a positive outcome may be evaluated better, thereby increasing this alternative's evaluative advantage over other alternatives. For example, imagine a decision maker wants to buy a house and out of the houses she can choose between, she prefers one that is close to her work place but quite expensive. If she engages in attractiveness restructuring, she may evaluate the high price as less negative and the fact that the house is close to her work place as more positive than initially. *Importance differentiation* means that the importance of attributes, as, for example, expressed in attribute weights is changed. The weight of an attribute on which an alternative has a negative outcome may be decreased and/or the weight of an attribute on which the alternative has a positive outcome may be increased. In the case of the house-buying decision, the decision maker may find price

[4] Note that Svenson's terminology is inconsistent. He uses the term 'differentiation' both specifically for pre-decisional processes and in a more general sense for all processes which increase the evaluative advantage of one alternative over the others. It is the latter meaning we refer to here.

less important and distance to work more important than before. Importance differentiation may be regarded as a particularly effective means of differentiation, because it allows enhancement of the preferred and devaluation of the non-preferred alternatives at the same time (Wernicke, 1993). When *facts restructuring* occurs, the individual changes his or her representation of known and/or newly emerging facts, often exploiting uncertainty associated with attribute outcomes. For example, the individual may start to doubt the validity of negative information about the preferred alternative or positive information about non-preferred alternatives. Also, attribute outcomes may be remembered in a biased way. After a while the decision maker may remember the price of the house she bought as lower than it actually was or the distance to her work place of the houses she did not buy as greater. Unless a decision is made over a long time period, memory biases would seem to be restricted to the consolidation phase, however. Finally, *problem restructuring* refers to changes in the representation of the decision problem, either by creating new attributes or by creating new alternatives. An example for the fact that the introduction of a new alternative which is dominated by one of the original alternatives increases the attractiveness and choice probability of the dominating alternative is the attraction effect (Huber, Payne, & Puto, 1982), as previously described in Chapter 3. Similarly, the introduction of a third alternative that dominates one of the initial two alternatives increases the attractiveness of the non-dominated alternative (Tyszka, 1983; Payne, 1982). Adding attributes on which the preferred alternative has good outcomes and the non-preferred alternatives have bad outcomes may be another way of increasing differentiation. If the decision maker, for example, brought to mind that her preferred house, apart from its other advantages, was also most likely to achieve the highest re-sell price, it would gain an additional advantage over the other alternatives. All these structural differentiation processes are assumed to operate in parallel and in conjunction with other differentiation processes, such as changes in criterion levels and the application of certain decision rules which favour the preferred alternative. For example, if a lexicographic decision rule is used, which prescribes choosing the alternative with the best outcome on the most important attribute, the most effective way for the decision maker to increase differentiation would be to concentrate any differentiation efforts on the most important attribute.

The structural differentiation processes suggested by Svenson can be related to the components of the MAUT equation (see Chapter 1), a fact that he fails to point out himself, however. Attractiveness restructuring is equivalent to changing $v(x_{ij})$, the evaluation of outcome i on attribute j, while facts restructuring means changing x_{ij}, the outcome for alternative i on an attribute j itself. Importance differentiation involves a change of w_j, the weight for an attribute j, and, finally, problem restructuring may change the number of attributes entered into the equation. This mapping of structural differentiation processes onto the MAUT linear model components allows one to estimate the extent of differentiation and consolidation that the decision maker exhibits by measuring MAUT components during and after the decision process. This strategy was used in the empirical work reported in Chapters 6 and 7.

Research conducted within the framework of Diff Con theory so far has mostly concentrated on post-decisional structural consolidation. Typically, participants are presented with two alternatives described on a number of attributes by means of ticks on attractiveness scales which are anchored with ‘very bad’ and ‘very good’. Before participants make a decision between the alternatives, they usually rate the importance of each attribute by placing a crossing line on a visual analogue scale. After a time interval which may last five minutes (Svenson & Benthorn, 1992) up to several months (Svenson & Shamoun, 1997), participants have to reproduce the attractiveness descriptions and their importance ratings. Results have supported the assumption that post-decisional consolidation occurs and usually focuses on the most and/or the second-most important attribute (e.g., Svenson & Benthorn, 1992; Svenson & Malmsten, 1991, 1996; Benthorn, 1994). Also, attractiveness restructuring is said to be more pronounced than importance differentiation (Malmsten, 1996). This claim should be viewed with caution, however. It may be expected that the type of consolidation that is preferred depends on the ease with which it can be performed and its effectiveness. Such variables have not been systematically manipulated in investigations so far. Also, importance differentiation seems to have always been measured after attractiveness restructuring. Participants may take the first opportunity which presents itself (i.e., attractiveness restructuring) to achieve a satisfactory level of consolidation, and therefore not show much importance differentiation (see Götz-Marchand, Götz, & Irle’s similar criticism of the selective exposure paradigm above).

Another criticism concerns the measurement procedure for attractiveness restructuring. Describing the attractiveness of alternatives rather than directly describing the alternatives' outcomes on these attributes means that an important step, namely the mapping of utilities onto outcomes, is taken by the researcher rather than the participant him- or herself. Participants are then asked to reproduce the attractiveness descriptions from memory. Any changes in attractiveness representations, therefore, strictly speaking, do not reflect attractiveness restructuring but facts restructuring and, at best, a mixture between the two. It may be argued that describing attribute outcomes and asking participants to (re-)assess the attractiveness of these outcomes provides a stronger test for changes in the decision maker's value system (Wernicke, 1993). With the procedure described above, it is likely that by the time post-decisional consolidation is measured, individuals are not able to remember much about the alternatives except the choice they made. They may then use their choice as a clue for reconstructing attractiveness descriptions (or their own importance judgements) in a way that favours this alternative. A similar reconstructive process has been suggested to contribute to hindsight bias (McCloskey & Zaragoza, 1985; Pohl, 1993). Consistent with this assumption, significant consolidation is usually only observed after some time has passed (Svenson & Benthorn, 1992). Also, instructing participants to memorise their decision (not the description of the alternatives) results in a higher level of consolidation (Svenson, Ortega Rayo, Andersen, Sandberg, & Svahlin, 1994). The visual description of alternatives by ticks on a line, which may be more difficult to memorise than numbers, favours the expected consolidation effect. Benson III & Svenson (1992) indeed reported that consolidation was not observed when participants had to rate the attractiveness difference between the chosen and non-chosen alternative on a numerical scale, but was observed when participants had to make judgements on visual analogue scales. Also, Wernicke (1993) failed to find attractiveness and fact restructuring as well as importance differentiation when describing attribute outcomes in numerical terms.

A pre-condition for consolidation seems to be the decision maker's sufficient involvement in the decision (Svenson, Ortega Rayo, Andersen, Sandberg, & Svahlin, 1994). Involvement is assumed to increase with the relevance of a decision and the decision maker's feeling of personal responsibility for its consequences. Svenson and Malmsten (1991, 1996) asked participants to make a consequential choice between tickets for two different lotteries in

which one of two portable cassette players could be won. Whether participants received the lottery ticket of their choice was determined by throwing a die. Immediate outcome feedback resulted in more and quicker consolidation for consequential than for non-consequential decisions; participants who won a ticket for their preferred lottery increased their perception of their prior chances of winning a ticket, whereas participants who did not win a ticket decreased their perception of their prior chances. Also, a successful outcome resulted in a higher degree of facts consolidation than an unsuccessful outcome. Svenson and Hill (1996) and Svenson and Shamoun (1997) investigated students' real decisions about which area of specialisation to choose for their final year of study. Participants' preferences and values were elicited at different stages of the decision process: prior to the decision, after the decision but before the implementation, and several months after the implementation. Svenson and Shamoun found that participants with a strong value conflict (i.e., those whose chosen alternative was worse than another on an important attribute) showed differentiation and consolidation that was strongly focused on the conflict attribute and was stronger than for participants without such a conflict.

Contrary to expectations, Svenson, Ortega Rayo, Andersen, Sandberg and Svahlin (1994) repeatedly found no facts restructuring consolidation, when participants were told that they would have to justify their decision afterwards. A possible reason for this may be the fact that participants were made accountable before they made a decision. As reported earlier, Tetlock and his co-workers (1985a, 1992; Tetlock, Skitka & Boettger, 1989) have shown many times that pre-decisional accountability leads to a deeper, more effortful information processing compared to no accountability. As a result of this deeper processing, participants in Svenson et al.'s (1994) experiment may have developed a better memory for the alternatives' descriptions and, hence, were less likely to show consolidation. In line with this explanation, the hindsight bias, which seems to be very similar to post-decisional consolidation in terms of its underlying processes, has been found to decrease as the amount of cognitive effort expended increases (Creyer & Ross, 1982).

The ideas expressed in Diff Con theory are not really new, its value lies in the fact that it has successfully managed to combine ideas from the social psychology and decision making literature, in particular dissonance theory and MAUT. The assumption of a multiattribute representation of decision alternatives allows one to test predictions of the

different ways in which the evaluative advantage of the chosen alternative may be increased more easily than the holistic representation suggested by dissonance theory. Also, unlike dissonance theory, Diff Con theory recognises that already pre-decisional information processing may be biased to support a tentatively chosen alternative. However, very few studies conducted within the framework of Diff Con theory have addressed this issue at all (Benson III, 1993, Svenson & Hill, 1996). Both dissonance theory (in its revised version) and Diff Con theory stress the importance of a certain commitment or involvement of the decision maker as a pre-condition for the occurrence of their postulated effects. This poses the danger for both theories to become impossible to falsify, since a failure to find the expected effects can always be explained by a lack of commitment or involvement on behalf of the decision maker. Diff Con theory also shares with dissonance theory the problem of being quite loose, especially in terms of predicting which particular way of increasing the evaluative advantage of the chosen alternative will be adopted. Dissonance theory has attempted to solve this problem by introducing the rather circular concept of resistance to change (the cognition with the smallest resistance to change will be altered; if a cognition has been changed, it must have had the smallest resistance to change). Svenson's outline of Diff Con theory so far has addressed the issue of strength of differentiation and consolidation (e.g., it is less strong, if the chosen alternative initially already appears much more attractive than its contenders than if the perceived attractiveness difference is small, because in this case part of the necessary differentiation has already been achieved), but does not say anything about what determines which particular differentiation and/or consolidation process will be chosen and clearly needs more specification in this respect. Finally, whereas dissonance theory is clearly a hot cognition, motivational theory, through postulating a drive to reduce dissonance, the 'temperature' of Diff Con theory (Arkes, 1996) is not readily apparent. Diff Con theory assumes that the decision maker wants to protect him- or herself against the threats of post-decisional preference changes, but does not go into much detail about the exact nature of this motivation. At the same time, it specifies processes that may be regarded as "by-products of a fully functional information processing system" (Arkes, 1996, p. 270). It seems therefore that Diff Con theory may best be regarded as a hybrid model, like Baumeister & Newman's (1994) theory of self-regulated inference or other such models, for example, by Kruglanski (1980, 1989, 1990; Kruglanski & Ajzen, 1983), Pyszczynski and Greenberg (1987) and Kunda (1990, 1999).

Summary and conclusions

In this chapter we presented both theoretical and empirical work relating to motivated information processing, in particular, processes of confirmatory information search, evaluation and integration. We first outlined Baumeister and Newman's (1994) theory of self-regulated inference and decision processes, which describes how a decision maker with either the goal of reaching an accurate conclusion (referred to as an intuitive scientist), or the goal of reaching a particular conclusion (referred to as an intuitive lawyer), may regulate his or her information search, evaluation and integration in order to achieve these respective goals. This was followed by a review of the empirical literature on various types of confirmation bias. In the domain of information search, confirmation bias was found to manifest itself either in the form of a hypothesis preservation bias, that is, the tendency to search for information that has a high likelihood of supporting one's hypothesis or preferred conclusion, or in the form of a positive test strategy, the tendency to test cases that are expected (or known) to be true under the working hypothesis, that is, to ask questions in which a 'yes' or positive answer would confirm one's hypothesis. The former has mainly been explored within the framework of cognitive dissonance (Festinger, 1957, 1964), the latter within Wason's (1960) rule discovery paradigm and variations thereof. In line with Baumeister and Newman's (1994) assumption that it may be difficult for an individual with the mindset of an intuitive lawyer to search information in a confirmatory manner (given that it is often difficult to know whether information will be supportive or not before looking at it), empirical results in this area of investigation are inconclusive. Where it comes to confirmatory information evaluation and integration, however, results are more clear-cut; there is abundant evidence that individuals bias their evaluation of information and integrate it in such a way that their preferred conclusion is supported. This may be done in several ways. Svenson's differentiation and consolidation theory specifies the processes that may be observed in multiattribute decision making, including re-assessing the attractiveness of attribute outcomes (attractiveness restructuring), falsely remembering the outcomes themselves (facts restructuring), changing attribute weights (importance differentiation) and changing the representation of the decision problem, either by creating new attributes or by creating new alternatives (problem

restructuring). Also, the decision maker may choose to apply certain decision rules which favour the preferred alternative.

In conclusion, the application of particular processing strategies appears highly adaptive and situation-specific in order to keep cognitive effort minimal, an observation that has been made many times in various research contexts, and has resulted in the proposal of various cost-benefit frameworks in the decision literature, such as Simon's (1955) concept of bounded rationality, Beach and Mitchell's (1978) cost-benefit strategy selection model, extended by Christensen-Szalanski (1978), and Payne, Bettman and Johnson's (1993) effort-accuracy framework.

All of the aforementioned effects may occur not only when there is a personal motivation to arrive at a certain conclusion but also when the individual just starts off with a certain expectation or hypothesis, rendering the distinction between so-called 'hot' (motivated) and 'cold' (non-motivated) cognition (Abelson, 1963) obsolete. In addition, although Baumeister and Newman's theory distinguishes between accuracy and directional goals, it would be wrong to assume that the individual will only ever pursue one or the other, and therefore either show thorough and objective or biased and selective information processing. Rather, any attempt at self-regulation of cognitive processes is always subject to reasonableness constraints (Kunda, 1990) and individuals attempt to maintain an illusion of accuracy and objectivity (Pyszczynski & Greenberg, 1987). This means that they are not free to arrive at any conclusion they want, but need to observe the constraints of the evidence available to them. Furthermore, directional goals will usually be paired with accuracy goals, and the individual has to deal with potential conflicts between the two. Accuracy goals have commonly been found to induce a deeper information processing, particularly a more extensive information search. Ironically, this more extensive processing caused by accuracy goals may *enhance* rather than reduce bias, by facilitating the construction of justifications for the desired conclusion (Kunda, 1990). This contradicts the belief expressed by many authors that strong accuracy goals will minimise the impact of any directional goals (e.g., Pyszczynski & Greenberg, 1987). One situation identified by Kunda (1990) as arousing *both* accuracy and directional goals is outcome dependency, that is, a situation in which individuals expect their own outcomes to depend in some way on another person, for example, because this person will be their partner in a later task and rewards will be determined by joint performance. From

what has been said about accountability so far, it should be obvious that any accountability manipulation which involves external accountability contains a strong element of outcome dependency, in that the individual expects to be evaluated by the people he or she is accountable to. It may therefore be expected that external accountability will make both accuracy and directional goals salient and will be associated with the cognitive effects characteristic of both. In the next chapter, we will propose a process model of accountability which will integrate these and earlier conclusions.

CHAPTER 5: A PROCESS MODEL OF ACCOUNTABILITY

This chapter will summarise the conclusions of previous chapters and integrate them into a process model of accountability (PMA in the following), which specifies how accountability demands may affect the processes of information search, evaluation and integration under various conditions of accountability. This model provides the basis for the empirical work undertaken, which will be outlined at the end.

A process model of accountability: Integration of previous conclusions

The discussion of motivational reasoning processes in the previous chapter implied that decision processes may be affected by two conflicting motives of the decision maker, accuracy goals, that is, a need to make an objective or normatively correct decision, and directional goals, that is, a need to arrive at a particular decision. Accuracy goals make the decision maker act like an intuitive scientist who conducts a thorough and unbiased information search, attempts to evaluate information objectively and integrates information according to rules and criteria set in advance, whereas directional goals prompt the decision maker to act like an intuitive lawyer who searches information selectively and biases information evaluation and integration in such a way that the preferred conclusion is favoured (Baumeister & Newman, 1994). In order to link accountability to these motives, we need to specify how accountability manipulations affect the salience of accuracy or directional goals.

The analysis of the accountability concept in Chapter 2 suggested that a context in which the decision maker is made externally accountable for his or her decisions will contain two types of cues, (1) responsibility cues that establish a causal link between the decision maker's choice and its consequences and suggest that the decision maker has control over the decision and thereby, at least partly, its associated consequences, and (2) evaluation cues that suggest that others are present who will evaluate the decision maker's choice with regard to certain norms or standards and imply that the decision maker's performance and /or conclusion is identifiable to them. Responsibility cues create a feeling of personal responsibility and focus the decision maker's attention on internal norms and values. The

PMA assumes that personal responsibility strengthens the decision maker's need to make a correct decision, that is, increases the salience of accuracy goals. Evaluation cues, on the other hand, create evaluation apprehension and increase the decision maker's focus on external norms and standards. Evaluation apprehension is assumed to make directional goals salient; it increases the decision maker's need to make a particular decision, namely one that is consistent with salient standards of evaluation and therefore can be easily justified to the people the decision maker is accountable to. The direction the decision process takes is suggested by the perceived norms and standards for evaluation. In some cases, the accountability manipulation itself may either explicitly provide or implicitly suggest such norms or standards, for example, by giving information about the preferences of the audience the decision maker has been made accountable to. In other cases, the decision maker may not be aware of either audience preferences or other external norms and will then rely on his or her personal norms and values which have been made salient by the responsibility cues inherent in the accountability manipulation.

The need to be accurate, as discussed in the previous chapter, is predicted to result in more complex information processing, in particular, a more extensive information search, because this part of the decision process can be most easily regulated towards objectivity. Directional goals, on the other hand, are assumed to primarily affect information evaluation, in such a way that the alternative which appears to be most justifiable or the alternative the person has already made a commitment to is favoured. This regulation is aimed at the resulting decision rather than the process by which the decision is arrived at. Information search and information evaluation are therefore seen as primary means to achieving different goals, namely accuracy on the one hand and being able to justify the preferred decision on the other. Information integration occupies an intermediate position in this respect. It can support accuracy goals by the application of complex decision rules that take into account all the information searched before, but it can also support directional goals by the application of decision rules that favour the preferred alternative. A graphical representation of the PMA predictions is shown in Figure 5.1.

A central assumption of the PMA is that accountability does not normally increase the salience of *either* accuracy *or* directional goals, but that the cues inherent in the accountability context will increase the strength of *both* these motives and, hence, an effect of accountability

on depth of information search and complexity of integration as well as degree of bias in information evaluation and integration will be observed. The seemingly contradictory results of previous research on the effects of accountability on decision processes, which suggest that accountability may both result in more complex and more biased information processing compared to a situation where no accountability pressures are present, are therefore well predicted by this model.

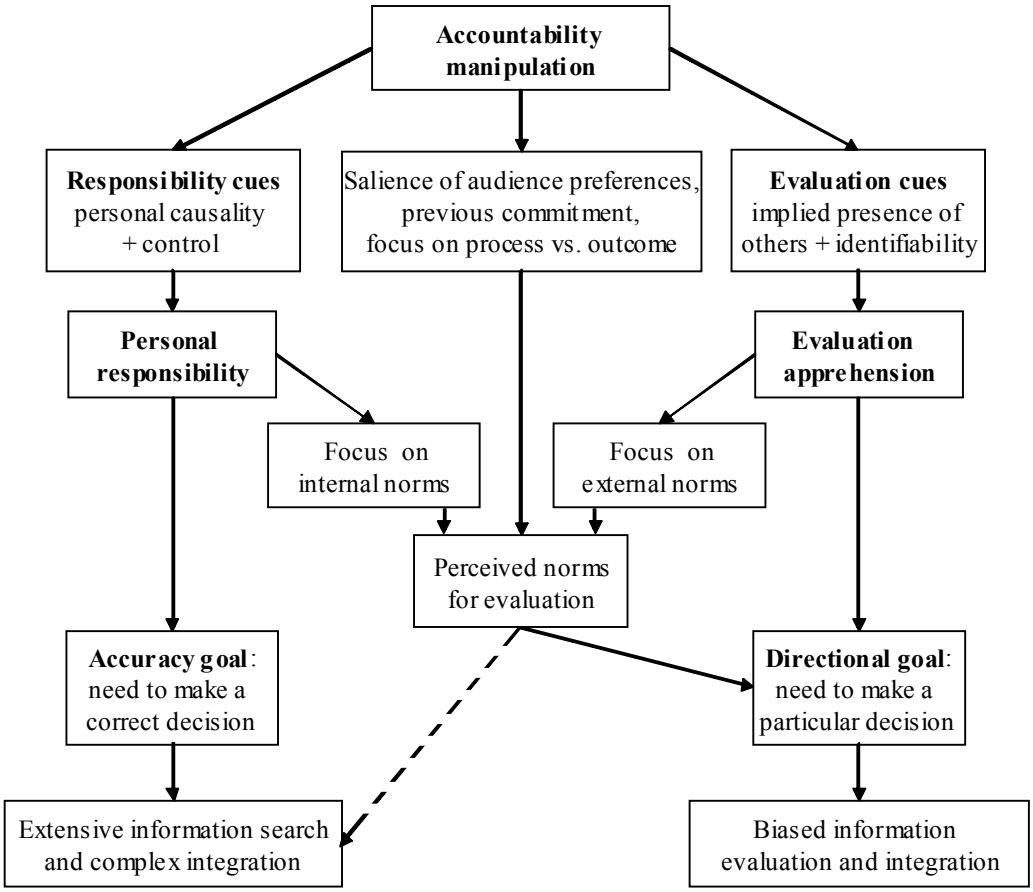


Figure 5.1. The process model of accountability (PMA)

The model also takes into account that the processes of information search, evaluation and integration do not operate independently of each other, but that they mutually reinforce each other (depicted by the double arrow in Figure 5.1), as discussed in the previous chapter. Information that has been searched by the decision maker in a first step may be used to bias the decision process at the evaluation stage. Likewise, a biased information evaluation may affect the type of information that is searched when the decision maker returns to the

information gathering stage, for example, because the previous information does not seem to be strong enough to support the desired decision.

Having introduced the basic assumptions of the PMA, the next step will be to outline how the effects of the different sub-types of accountability discussed in Chapter 3, namely, accountability to an audience with known vs. unknown views, pre- vs. post-decisional accountability and procedural vs. outcome accountability may be explained by the PMA.

According to Tetlock's social contingency model (1985a, 1991, 1992), pre-decisional accountability results in complex, compensatory information processing if the preferences of the audience the decision maker is accountable to are not known, whereas post-decisional accountability to an audience with unknown views, that is, being made accountable after one has already made a commitment to one of the alternatives, results in biased, defensive information processing. The PMA can explain these effects by assuming that the responsibility cues inherent in pre-decisional accountability to an audience with unknown views will particularly make accuracy goals salient, while the salience of any directional goals will be relatively low, due to the absence of perceived audience preferences or other external norms or standards which could drive the decision in a particular direction. This accountability manipulation will therefore make information processing more complex, and the decision maker is expected to make a decision that closely matches his or her personal preferences, as these are the only standards available to him or her. Having previously made a commitment to a particular alternative, however, will provide the decision maker with a strong directional goal, namely to defend the appropriateness of his or her previous decision. At the same time, the responsibility cues still inherent in this manipulation are predicted to induce a more extensive information search, compared to a situation of no accountability. It is expected, however, that the higher depth of processing serves the directional goal of defending the decision maker's past decision. Note that the prediction of a more extensive information search whenever responsibility cues are present, no matter whether accountability is pre- or post-decisional, departs from Tetlock's predictions and, hence, provides a suitable way of testing the two models against each other. In fact, the results of Tetlock's own research point to the correctness of this PMA hypothesis. Tetlock, Skitka, and Boettger (1989), for example, found that introducing post-decisional accountability demands was not sufficient to produce simplification of thoughts compared to pre-decisional accountability.

Tetlock (1985a, 1991, 1992) furthermore states that decision makers apply a low-effort acceptability heuristic and match their decision to the preferences of the audience they are accountable to if those are known or assumed, but not if they have already committed themselves to a different decision. The PMA arrives at the same prediction, assuming that while both the knowledge of or assumptions about audience preferences and the decision maker's past commitment make certain directional goals salient, past commitments do so more powerfully, either owing to implications for the decision maker's self-image or owing to the greater cognitive availability of previously processed information. It may also be expected that the decision maker is more likely to adjust his or her expressed preferences to those of the audience if his or her own preferences are similar to the presumed audience preferences, because in such a situation there is little conflict between different directional goals.

Finally, the fact that procedural accountability typically results in more complex information processing, whereas outcome accountability results in more biased information processing can be explained by the PMA in the following way. Procedural accountability signals to the decision maker that the standards against which his or her decision will be measured relate to the decision process rather than the decision outcome. This suggests that a thorough processing of all available information is a suitable way of meeting these standards and avoiding blame (see Siegel-Jacobs & Yates, 1996, for a similar argument), and, hence, information search will be deep and information integration complex. Outcome accountability, on the other hand, signals to the decision maker that his or her decision will be considered acceptable if it results in a successful outcome. Given that in such a situation it is impossible for the decision maker to totally safeguard him- or herself from failure by processing information very carefully, he or she will try to adopt strategies that maximise the likelihood of a successful outcome, for example, by attending to information about what the evaluators would do in a similar situation. It is therefore expected that under outcome accountability, the decision maker will not invest particular effort into the decision process but will simply make a decision that is consistent with the preferences of the audience he or she is accountable to, if these preferences can be inferred or have been made salient in any way.

Summary and research plan

The PMA, in departure from previous models of accountability effects on judgement and decision processes, suggests that any accountability manipulation will usually create both accuracy and directional goals in the decision maker, and as a consequence, will induce information processing that is both more complex and more biased compared to a situation without accountability pressures. The exact nature of the accountability manipulation, however, will determine the relative strength of these two effects. A manipulation that primarily contains personal responsibility cues but minimises evaluation cues, as it is typical when trying to create internal accountability, is expected to mainly result in deeper and more complex information processing while any biasing effects will be weak. On the other hand, a manipulation that makes evaluation cues particularly salient, including audience preferences or previous commitments that may serve as standards of evaluation, is expected to result in both complex and strongly biased information processing, as long as the available evidence permits a biased conclusion. A test of the validity of the PMA should therefore involve manipulating personal responsibility and evaluation cues as well as perceived norms for evaluation and observing the effects of these manipulations on the complexity and degree of bias in information search, evaluation and integration.

The first study to be reported here investigated all three of these factors in an experiment testing the effects of internal vs. external accountability and pre- vs. post-decisional accountability on the decision process. Internal accountability was manipulated by creating a scenario in which participants were made aware of their personal responsibility for the consequences of their decision, while the decision itself was said to be kept anonymous to others, thereby lowering evaluation concerns. External accountability was manipulated in a similar way, except that participants were explicitly told that others would be aware of their decision, thereby strengthening evaluation apprehension. There was no explicit manipulation of audience preferences. Instead, the assumed influence of perceived norms of evaluation was tested by comparing the effects of pre-and post-decisional accountability, predicting that post-decisional accountability would bias the decision process towards supporting a previous decision.

The second study that tested the PMA focused on the effects of external accountability to different interest groups whose preferences could be easily inferred by the participants and were therefore expected to direct the decision process towards a particular conclusion. The data were also analysed with respect to the absence or presence of conflict of audience preferences with the decision maker's own preferences. Again, the timing of the accountability manipulation was varied, in order to test the relative strength of bias induced by audience preferences vs. bias induced by a previous commitment. The results of these studies will be reported in Chapters 6 and 7.

CHAPTER 6: THE EFFECTS OF INTERNAL VS. EXTERNAL ACCOUNTABILITY

Introduction

The first experiment to test the validity of the process model of accountability (PMA) presented in Chapter 5 addressed effects of internal and external accountability on information search, evaluation and integration. Although a distinction between these two types of accountability has commonly been made in the literature, studies *directly* investigating the differential effects of these two types of accountability have not been reported in the literature so far. According to the PMA, internal accountability is assumed to arise when individuals feel personally responsible for their decision, that is, when they feel that the decision rests on them and they are therefore responsible for its consequences, while there is no opportunity for external evaluation of their decision, because their contribution is not identifiable. This means that impression management concerns and evaluation apprehension are minimal and the biasing effects of accountability should therefore not be as strong as in a situation of external accountability, where decision makers perceive a personal responsibility for their decision and, more importantly, their performance can be evaluated by other people, because it is identifiable.

According to the PMA, there will be no difference between these two types of accountability in terms of depth of information search and complexity of information integration when making an initial decision. This is, because in both cases responsibility cues are predicted to make accuracy goals salient which, in turn, will result in an extensive information search and complex, compensatory information integration. There will, however, be a difference between internal and external accountability with regard to the degree of bias evident in information evaluation and integration. Externally accountable individuals, because of the fact that they anticipate their performance to be evaluated by other people, will look for an alternative that they can easily justify to those to whom they are accountable. If a norm has been made salient against which the decision maker's performance will be evaluated, for

example, through suggesting that the person or group the decision maker is accountable to has a preference for a certain alternative, adjustment to this norm can be expected (Tetlock, 1990, 1992, Tetlock, Skitka, & Boettger, 1989). If a norm has not been made salient, there are two possible reactions. Firstly, the decision maker may try to infer the norm and then adjust his or her information evaluation and integration to the inferred norm. Secondly, the decision maker may be led by internal norms and opt for the alternative with the best outcome on the attribute(s) which appear(s) most important to him or her in the given decision context, that is, make a choice that is consistent with a lexicographic decision rule.

Another way of enhancing the justifiability of an alternative, which has been suggested by Diff Con theory (Svenson, 1992, 1996), is to increase the evaluative difference between the chosen alternative and the non-chosen alternatives. This may be achieved by structural differentiation processes, such as attractiveness restructuring or importance differentiation. Attractiveness restructuring refers to a better evaluation of attribute outcomes for the chosen alternative and/or a worse evaluation of attribute outcomes for the non-chosen alternatives. Importance differentiation refers to increasing the perceived importance of attributes on which the chosen alternative performs well and/or decreasing the perceived importance of attributes on which it performs badly. These processes should also find their expression in the holistic evaluation of alternatives, that is, the difference in judgements of overall attractiveness between the chosen alternative and the non-chosen alternatives should be enlarged. Such differentiation processes should be stronger when participants are made externally accountable than when they are made internally accountable.

The difference between the effects of internal and external accountability should be diminished, however, when individuals are only made accountable once they have already committed themselves to an alternative. In this case, both internally and externally accountable individuals are expected to bias their information search as well as their information evaluation and integration so as to bolster their previously chosen alternative. To what extent this tendency will be stronger for participants under internal or external accountability may be expected to depend on whether participants who are made externally accountable perceive their choice to be in conflict with the preferences of the audience they are made accountable to. Although they will feel obliged to search information about the alternative their audience seems to prefer, resulting in a less biased information search

compared to that of internally accountable participants, externally accountable participants will bias their information evaluation and integration more strongly than internally accountable participants, because of an increased need to be able to justify that their previously made decision was a good one. If there is no conflict between externally accountable participants' choice and the perceived preferences of their audience, on the other hand, both internally and externally accountable participants may be expected to show a similar degrees of bias towards supporting their chosen alternative, once they have been made accountable.

The experiment conducted to test the predictions of the PMA therefore manipulated the following two factors: type of accountability (internal vs. external) and time when accountability was introduced (before a first decision vs. after a first and before a final decision). Participants were made internally or externally accountable either already before they made a preliminary or only before they made a final decision about which one of four job applicants to select. Information search was assessed by presenting participants with an information board (Payne, 1976), which they used to search information about the four different applicants who had been described on five attributes. Participants had to search information twice, once before they made their first decision, and a second time before they made their final decision. A particular feature of the two information sets used was that the attributes and attribute outcomes for the second search were largely redundant with those of the first search. For example, if an attribute in the first set was whether or not the applicant had had an educational stay in a foreign country, a redundant attribute in the second set would be the applicant's knowledge of foreign languages. This was to enable participants to foresee the value of the information they searched and so make it possible to test whether information search after a commitment was biased to support the previously chosen alternative. Information evaluation and integration was measured by participants' choices, their overall judgements of the suitability of each candidate for the job, and their evaluations of attribute outcomes and attribute weights, which were elicited before participants made their first decision and after they made their final decision.

For the *first decision,* it was expected that *information search* would be more extensive, less variable across alternatives and attributes, more compensatory and less

strongly concentrated on the chosen alternative in conditions with either internal or external accountability compared to conditions with no accountability.

The effects of the manipulated variables on the *first information evaluation and integration* were expected to be the following. It was predicted that, compared to participants who were not or were internally accountable, participants who were externally accountable would be more likely to choose the alternative with the best outcome on what they regarded the most important attribute. This was, because no particular audience preferences were made salient and a choice consistent with such a lexicographic decision rule is easy to justify (see Slovic, 1975). It was also predicted that externally accountable participants would show a larger difference between their suitability judgement for the chosen alternative and their suitability judgement for the average non-chosen alternative than participants who were not accountable or internally accountable. Since unidimensional values and weights were only elicited before the first and after the final decision, attractiveness restructuring and importance differentiation could not be tested at this point.

For the *final decision*, the PMA predicted that the *information search* of participants who had been made either internally or externally accountable immediately beforehand would display characteristics of depth and compensation more strongly than that of participants who had already been made accountable before the first information search. This is, because for participants who had been made accountable immediately prior to this information search, the invoked accuracy goals should have been more salient than for participants who had already been made accountable before the first decision.

For participants for whom accountability was introduced only after they had already made a first decision, it was furthermore expected that, although their information search would be more extensive in order to meet salient accuracy goals, it would also be more biased to support the previously chosen alternative, compared to that of participants who had been made accountable from the start.

The *final information evaluation and integration* was expected to be generally biased in the direction of favouring the previously chosen alternative, because all participants would be motivated to confirm their previously made decision. However, this motivation should be strongest for participants who had been made externally accountable immediately prior to the

final decision and weakest for participants who had been made internally accountable before the first decision, with the other two conditions in-between. As an expression of biased information evaluation, participants were expected to increase the evaluative difference between their chosen and average non-chosen alternative. This was predicted to become apparent not only in participants' overall evaluations of the job applicants but also in a change of unidimensional values and weights, as suggested by Diff Con theory (Svenson, 1992). Specifically, it was expected that negative attribute outcomes of the chosen alternative would be evaluated less negatively and/or positive attribute outcomes would be evaluated more positively than before a decision had been made. It was also expected that the weights of attributes on which the chosen alternative had positive outcomes were increased, and/or the weights on which the chosen alternative had negative outcomes were decreased.

Method

Participants

Seventy-two psychology students at the Technical University of Darmstadt participated in the experiment, 66 women and 6 men. Their age ranged from 20 to 33 years (M = 25.5), the number of semesters they had been studying psychology varied between 2 and 14 (M = 4.8). Participation was voluntary. Instead of a fixed reward, participants received a lottery ticket with which they could win up to DM 500. After completion of the data collection, 13 winners were drawn.

Design

The design was a 2 x 2 factorial design, manipulating type of accountability (internal vs. external) and time of introducing accountability (before a first decision vs. before the final decision). Both factors were varied between subjects, resulting in four experimental groups with 18 participants each.

Materials

The decision task presented to the participants was one of personnel selection. In order to create a realistic scenario student participants could relate to and would be involved in, the decision situation chosen was the selection of a junior lecturer (the post of Wissenschaftlicher Mitarbeiter at German universities). In order to be able to test the extent to which information search was biased after a commitment to an alternative had been made, two sets of highly redundant information had to be constructed. To find the attributes for these information sets, a pilot study was conducted. Sixty psychology students at the Technical University of Darmstadt (none of which participated in the main experiment) rated 14 qualification criteria for the job of junior lecturer in Organisational Psychology. These criteria had to be rated with regard to their importance for selecting an applicant (on a scale from 1 'not at all important' to 5 'very important') and all possible pairs of these criteria[1] with regard to their similarity (on a scale from 1 'very dissimilar' to 5 'very similar'). Ten of the 14 rated criteria were then selected to be used as attributes in the main experiment, five for each information set. The information sets were created in such a way that the five attributes in each set were judged maximally dissimilar from each other (i.e., were perceived to describe clearly different aspects of the qualification for the job), while at the same time each attribute from the first information set was judged as highly similar to one (and only one) attribute from the second information set. Pairs of highly similar attributes will in the following be referred to as parallel attributes. The resulting information sets, with mean similarity and importance ratings for their attributes, are shown in Table 6.1.

For each of the five attributes in information set 1, four different outcomes were specified, which were spread equally across a realistic range of outcomes. The attribute outcomes were used to construct four alternatives, that is, job candidates. To make the decision between them difficult, the overall quality of the alternatives had to be similar. This was achieved by assigning outcomes to the alternatives in such a way that their overall

[1] 91 different pairs of criteria were rated in total. Each person rated a selection of 49 pairs, 7 pairs for which high similarity was expected (resulting in 60 ratings per pair) and a selection of 42 of the remaining 84 pairs (resulting in 30 ratings per pair).

evaluation, predicted by an additive MAUT model (see Equation 1.1), was as similar as possible.

Table 6.1. Information sets used in the main experiment, with mean similarity and importance ratings.

Information Set 1			Information Set 2	
Attribute Labels	*Importance*	*Similarity*	*Importance*	Attribute Labels
educational stay in a foreign country	2.47	4.64	3.50	knowledge of foreign languages
activities as a student member of a professional association	2.72	4.03	2.53	activities as a student representative on university committees
diploma thesis grade	3.55	3.95	3.25	reference from thesis supervisor
number of placements in the area of Organisational Psychology	3.78	3.78	3.25	relevance of thesis topic to Organisational Psychology
job as a student research assistant	3.22	3.18	3.87	computing skills

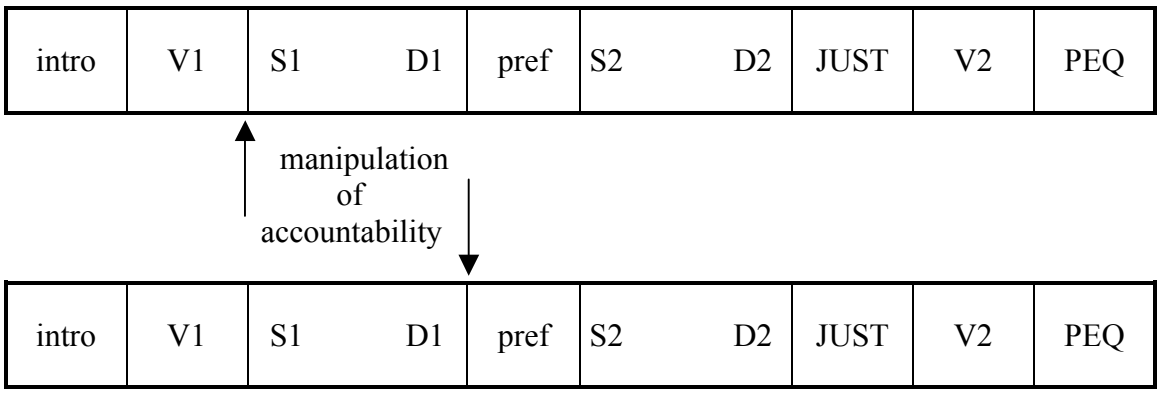
Note: Importance ratings were given on a scale from 1 (very unimportant) to 5 (very important), similarity ratings were given on a scale from 1 (very dissimilar) to 5 (very similar).

The mean importance judgements obtained in the pilot study were normalised and served as attribute weights in the model equation. Unidimensional values for the model equation were estimated by assuming a value of 0 for the worst of the four outcomes for each attribute, a value of 100 for the best outcome, and values at equal distances in-between for the other two outcomes (i.e., 33.33 and 66.67). Outcomes for the attributes of information set 2 were specified in such a way that there was a perfect correspondence between parallel attributes. This meant that the alternative with the best, second-best, etc. outcome on an attribute of information set 1 also had the best, second-best, etc. outcome on the parallel attribute of information set 2.

Procedure

Participants were tested in individual sessions. They received written instructions, which were handed to them by an experimenter who was present throughout the experiment and gave additional explanations if necessary. A schematic representation of the sequence of events during the experiment is shown in Table 6.2. In the following, each event will be described in detail:

Table 6.2. Procedure.



Note: Capital letters indicate elicitation parts, lower-case letters indicate instruction parts.
intro: introduction to the decision problem
V1, V2: first and second elicitation of unidimensional values and weights
S1, S2: first and second information search
D1, D2: first and second decision
pref: information about preferences of other committee members
JUST: justification of the decision
PEQ: post-experimental questionnaire

Intro: Introduction to the decision problem

Participants were told that their task would be to make a personnel decision. They were informed about the normal selection procedure at their department, namely that personnel decisions were made by a selection committee and that one of the committee members was always a student. Participants were asked to imagine being the student member of a such a committee, with the task to select a junior lecturer in Organisational Psychology. The committee was said to have five members; apart from the participant him- or herself, there were said to be two professors and two other junior lecturers on the committee. Participants were told about the duties the advertised job involved, and that four candidates had been short-listed and interviewed. For each of the four candidates, information was said to be available with regard to five different qualification criteria. The attributes of information set 1 were listed.

V1: First elicitation of unidimensional values and weights

In order to elicit unidimensional value functions, the attributes of information set 1 were presented, each with the four outcomes that had been previously specified. Each outcome had to be rated on a scale from 0 (very bad) to 100 (very good). The order of presentation of attributes as well as the order of presentation of attribute outcomes was randomised for each participant.

Weights for the attributes of information set 1 were elicited with the swing method (von Winterfeldt & Edwards, 1986). Participants were presented with those outcomes of each attribute they had rated worst and best. If a participant had rated all attribute outcomes identically, the objectively worst and best outcomes were used. Participants were asked to imagine a candidate with worst outcomes on all attributes and indicate the attribute whose 'swing' from the worst to the best outcome would result in the largest, second-largest, etc., improvement of this candidate's qualification, thereby effectively rank-ordering the attributes from most to least important. The swing resulting in the largest improvement was arbitrarily set to 100 and all other improvements had to be specified as percentages thereof. Final weights were obtained by normalising these percentages such that the weights added up to 1.

Manipulations of accountability

For half of the participants accountability was introduced at this point, the other participants did not receive any instructions and continued directly with the information search procedure. Half of the participants who received accountability instructions were made internally accountable, the other half were made externally accountable.

Internal accountability. Participants were informed that, in order to arrive at a committee decision, there would be a vote by ballot. The votes of single committee members would be kept secret (no identifiability). Participants were furthermore told that conversations with other committee members had suggested that a stalemate might be expected between their choices, so that their own vote would be decisive (personal responsibility).

External accountability. Participants received the information that, in order to reach a committee decision, there would be a public vote and everybody on the committee would know how each committee member had voted (identifiability). As in the 'internal accountability' conditions, participants were informed about the expected stalemate between the other committee members and that their own vote would be decisive (personal responsibility).

S1: Information search set 1

In order to make a decision, participants were asked to search information about the four short-listed candidates. Information search was monitored by employing the information board technique (Payne, 1976). Information about the candidates, named A,B,C, and D, was presented on cards arranged in an attribute-by-alternative matrix. On each card information was given about a particular candidate's (i.e., alternative's)

outcome on a particular attribute. At the beginning only attribute and alternative labels were visible, all other cards were turned over so that the information on them was not visible. Participants had to search information by turning over one card, reading the information provided on it, and then putting it back again so that the information was not visible anymore, before turning over the next card. This procedure was continued until participants were ready to make a decision. Participants were free to search as many items of information as they wanted to and were allowed to look at the same item as often as they liked. The sequence in which cards were turned over was recorded. Appendix 1 shows a schematic representation of the information board used for the first information search.

D1: First decision and suitability judgements

Participants indicated their chosen candidate and rated each candidate in terms of suitability for the job. This was done on a scale ranging from 0 (extremely unsuited) to 100 (extremely suited). Additionally, participants rated on 5-point scales how difficult the decision had been (1 = 'not at all difficult', 5 = 'very difficult') and how sure they were about having made a good decision (1 = 'not all sure', 5 = 'very sure').

Manipulations of accountability

Participants who had not been made accountable before now received either internal or external accountability instructions as described above. The remaining half of participants did not receive any instructions at this point.

pref: Preference of other committee members

All participants were told that a stalemate had arisen between the other committee members. Two committee members were said to have voted for the candidate the participant had chosen, whereas the remaining two committee members allegedly had voted for a different candidate. This other candidate was randomly chosen from the remaining candidates, with the constraint that the number of particular pairs of candidates presented to participants was balanced as much as possible across conditions. Participants were also told that, because the decision proved difficult, the committee had decided to draw upon additional information and that before they were expected to make a final decision about which candidate they preferred, they were going to be given the opportunity to search additional information about the candidates.

S2 Information search set 2

Participants were presented with an information board whose upper half contained the information items of information set 1 and whose lower half the information items of set 2. Those information items of set 1 that participants had looked at during their first search were permanently visible now, the others were not accessible anymore. Information items from set 2 could be searched like before. The preferred candidates were marked on the board. Appendix 2 shows the items of information set 2, in the way they were arranged in the lower half of the second information board.

D2: Final decision and suitability judgements

When they felt ready, participants again indicated their chosen candidate and rated each candidate in terms of suitability for the job, like before. They also responded again to the questions about how difficult the decision had been and how sure they were about having made a good decision.

JUST: Justification

Participants had to justify their choice in writing. Participants in the 'internal accountability' conditions were only asked to write down the reasons for their choice. Participants in the 'external accountability' conditions were told that the other committee members wanted them to give the reasons for their choice and that they should therefore write them down for them.

V2: Second elicitation of unidimensional values and weights

The elicitation procedure of V1 was repeated. Participants rated the outcomes of and specified weights for the attributes of information set 1.

PEQ: Post-experimental questionnaire

Participants were asked how much they had expected to have to justify their decision, how strongly they had had the wish to make a good decision, and how strongly they had felt a need to be able to justify their decision to themselves and to the other committee members at the first as well as the final decision. Ratings were made on 5-point scales, ranging from 'not at all' to 'very much'. Additionally, participants had to rate the attributes of set 2 with regard how important they were for the decision, on a scale ranging from 1 (not at all important) to 5 (very important) and they were asked to give similarity ratings for all 25 possible pairs comparing one attribute of set 1 with one attribute of set 2. Similarity ratings were given on scales ranging from 1 (very dissimilar) to 5 (very similar).

At the end of the experiment, participants were thanked for their participation, received their lottery ticket, and were asked not to talk about the experiment with other students.

Results

First decision

Information search

The records of the first information search were used to derive the following measures:

- Depth of search

  Measures of depth of search included both the percentage of total number of items searched and the percentage of different cells searched. The former may exceed 100% if participants search cells repeatedly.

- Variability of search across alternatives and attributes

  Variability of search is defined as the standard deviation of the number of items searched across alternatives or across attributes. Variability of search *across alternatives* is 0, if the same number of items is searched for each alternative, and is > 0, if a different number of items are searched for each alternative. Variability of search *across attributes* indicates the extent to which different attributes receive different amounts of search. It is 0, if the same number of alternatives is searched for each attribute, and is > 0, if a different number of alternatives are searched for each attribute. Although the two measures are obviously not completely independent of each other, they are not completely redundant either, as Klayman (1983) has pointed out. The two indices may be used to determine where a high variability of search comes from, that is, from not looking at all attributes or not looking at all alternatives.

  Variability of search across alternatives, in particular, has been linked to the type of decision strategy employed. A low variability of search across alternatives is consistent with a compensatory strategy, where the decision maker trades off attributes against each other and therefore searches a similar number of attributes for each alternative. If variability of search across alternatives is high, on the other hand, the decision maker is assumed to have employed a non-compensatory strategy, where alternatives may be

excluded from further consideration once they are found not to meet certain criteria, and therefore the number of attributes searched for each alternative varies. A similar argument may be made for variability of search across attributes, with a high variability of search across attributes indicating that certain alternatives have been eliminated from the set of possible choices.

Variability of search across alternatives and across attributes was computed both for the total number of items searched and for the number of different cells searched. Variability of search for the total number of items searched takes repeated searches of the same cell into account.

- Compensatory processing

  Koele and Westenberg (1995) argue that, in addition to variability of search across alternatives, depth of search should be considered when trying to infer the kind of decision strategy used. If all attributes are relevant for making a decision, compensatory decision strategies require the processing of a large proportion of the available information (high depth of search) as well as the processing of a relatively constant proportion of information per alternative (low variability of search across alternatives). Following this line of reasoning, Koele and Westenberg's index combines depth of search with variability of search across alternatives in a multiplicatory function to arrive at a measure of compensatory information processing. The index is defined as

$$C = \frac{N_{TS}}{N_{Max}} \cdot (1 - \frac{var}{var_{Max}}) \qquad \text{(Eq. 6.1)}$$

N_{TS}:	total number of information items searched
N_{Max}:	maximum number of information items that can be searched
var:	variability of search across alternatives
var_{Max}:	maximum variability of search across alternatives.

When there is an even number of alternatives, as it was the case here, and variability of search across alternatives is computed for the proportion of information searched per alternative, it can be shown that the maximum variability of search equals 0.5. Equation 6.1 can then be written as

$$C = \frac{N_{TS}}{N_{Max}} \cdot (1 - \frac{var}{0.5}) \qquad \text{(Eq. 6.2)}$$

The resulting values lie between 0 and 1. Higher values indicate more compensatory information processing.

- Pattern of search

  Pattern of search refers to the sequence in which information is searched. It can be either alternative-wise or attribute-wise. An alternative-wise search means that individuals search information within a particular alternative, that is, across different attributes, and then move on to do the same for the next alternative. It is usually associated with compensatory information processing, since it is consistent with trading off attributes within each alternative. An attribute-wise search is a search within attributes and across different alternatives. It is usually associated with non-compensatory information processing, since it is consistent with a strategy where alternatives are only considered further if their attribute outcomes fulfil certain criteria.

  In order to determine the search pattern, the number of alternative-wise and attribute-wise search transitions has to be established. This is done comparing the *n*th item of information searched with the *n*th + 1 item of information searched. An alternative-wise transition is observed, if the *n*th + 1 item of information searched is within the same alternative but a for a different attribute than the *n*th item. An attribute-wise transition is observed, on the other hand, if the *n*th + 1 item of information searched is within the same attribute but for a different alternative. Shifts occur if the *n*th + 1 item of information searched is neither for the same alternative nor the same attribute. They are, however, not considered in the search pattern index developed by Payne (1976). The Payne Index is defined as

$$PI = \frac{N_{TAlt} - N_{TAtt}}{N_{TAlt} + N_{TAtt}} \qquad \text{(Eq. 6.3)}$$

N_{TAlt}: Number of alternative-wise transitions

N_{TAtt}: Number of attribute-wise transitions

A value of -1 indicates a strictly attribute-wise search, a value of +1 indicates a strictly alternative-wise search. As Stokmans (1991) points out, the expected value is only 0, if the information board matrix contains an equal number of alternatives and attributes. In the case of an information board with four alternatives and five attributes, as it was used here, it can be shown that the expected value is 0.05. This means that in order to check whether information search followed a pattern that differed significantly from a chance pattern, the empirically found Payne Index had to be tested against a value of 0.05.

- Concentration of search

The concentration of search (CS) measure (Johnson, Meyer & Ghose, 1989) indicates the extent to which search is concentrated on the chosen alternative. It is defined as

$$CS = \left(\frac{N_{CAlt} \cdot N_{Alt}}{N_{TS}}\right) - 1 \qquad \text{(Eq. 6.4)}$$

N_{CAlt}: number of information items searched for the chosen alternative

N_{Alt}: number of available alternatives

N_{TS}: total number of information items searched

A value of 0 indicates that, on average, the same number of items was searched for all alternatives, that is, that there was no concentration of search on the chosen alternative. A value > 0 results if search was concentrated on the chosen alternative and a value < 0 results if search was concentrated on the non-chosen alternatives. Concentration of search was calculated both for the total number of items searched and different cells searched.

At the time of the first information search, two groups of participants had not received any accountability instructions yet. Since independent sample t-tests revealed no significant differences between them, they were pooled and constituted the ‘no accountability’ condition. Table 6.3 shows a summary of the results.

Table 6.3. Process measures for the first information search, as a function of type of accountability.

	Type of Accountability			
	No[1]	Internal[2]	External[2]	*F* (2, 69), *p*
1. Depth of search (in %)				
total # of items	191	206	176	< 1, *ns*
different cells	91	93	90	< 1, *ns*
2. Variability of search across alternatives				
total # of items	2.10	2.01	1.56	< 1, *ns*
different cells	0.38	0.30	0.23	< 1, *ns*
across attributes				
total # of items	2.70	2.42	2.27	< 1, *ns*
different cells	0.40	0.33	0.42	< 1, *ns*
3. Compensatory processing	0.79	0.84	0.83	< 1, *ns*
4. Pattern of search (Payne Index)	0.20	0.08	0.08	< 1, ns
5. Concentration of search				
total # of items	0.11	-0.02	0.04	1.89, .159
different cells	0.09	0.03	0.08	< 1, *ns*

[1] n=36
[2] n=18

1. Depth of search

Information search was generally very extensive. The mean percentage of total number of information items searched across all conditions was 191%, that is, on average almost twice as many items were searched as there were different items available. The mean percentage of different cells searched (M = 91%) showed that, on average, almost all of the available items (namely 18.28 out of 20) were searched at least once. Participants obviously took the task seriously and carefully looked and looked again at information before making a decision.

It had been expected that accountability, compared to no accountability, would lead to a more extensive information search. Independent sample t-tests comparing these two conditions were employed to test this hypothesis. Both for the percentage of total number of

items searched and different cells searched, no significant difference was revealed, the obtained values were virtually identical (total number of items searched: accountability: *M* = 191%, no accountability: *M* = 191%, different cells searched: accountability: *M* = 92%, no accountability: *M* = 91%).

In addition, univariate Analyses of Variance (ANOVAs) were carried out in order to test the effect of type of accountability (no vs. internal vs. external) on the percentage of total number of items searched and different cells searched. No significant effects were obtained.

2. Variability of search

The overall mean variability of search across alternatives and attributes, calculated for search of different cells, was relatively low (*M* = 0.32 across alternatives, *M* = 0.39 across attributes). This was not surprising, given that 58% of the participants searched all cells and therefore received a value of 0 for both indices. A more informative measure of variability of search when information search is exhaustive, is variability of search computed for the total number of items searched, because it also takes into account repeated searches of the same cells. The overall mean variability of search across alternatives computed for the total number of items searched was relatively high (*M* = 1.94), and the overall mean variability of search across attributes even higher (*M* = 2.52). For a random search of an asymmetrical matrix with more attributes than alternatives, the opposite would be expected. This pattern of results shows that certain information items received more attention compared to others when the same information was searched more than once, and that attributes were attended to more unevenly than alternatives.

It had been predicted that accountable participants would show a lower variability of search across alternatives and attributes than participants who had not been made accountable. The independent t-tests performed to test these hypotheses did not yield any significant effects, although the obtained differences were in the right direction (variability of search across alternatives: accountability: *M* = 1.78, no accountability: *M* = 2.10 for the total number of items searched, accountability: *M* = 0.26, no accountability: *M* = 0.38 for different cells searched; variability of search across attributes: accountability: *M* = 2.34, no accountability:

$M = 2.70$ for the total number of items searched, accountability: $M = 0.37$, no accountability: $M = 0.40$ for different cells searched).

The one-factorial ANOVAs carried out to test the effect of type of accountability also showed no significant effects.

3. Compensatory processing

Consistent with the results for depth of search and variability of search across alternatives for different cells, the overall mean of Koele and Westenberg's Index of Compensatory Processing, which combines these two measures, was high ($M = 0.82$), suggesting that information processing was highly compensatory.

The predicted significant difference in the extent of compensatory processing between accountable and non-accountable participants was not found, although, as expected, information search tended to be more compensatory for accountable than for non-accountable participants ($M = 0.84$ vs. $M = 0.79$).

There was also no significant main effect of type of accountability on the index of compensatory processing.

4. Pattern of search (Payne Index)

The mean Payne Index across all conditions was 0.14, suggesting that search was slightly more alternative-wise than attribute-wise, although the difference to the expected value of 0.05 for a balanced search was not significant ($t(71) = 1.50$, $p = 0.138$).

It had been expected that accountable participants would be significantly less likely to show an alternative-wise search than non-accountable participants, suggesting that they had processed information in a compensatory manner. The observed difference between these two conditions pointed to such an effect, in that accountable participants tended to show a less alternative-wise search than non-accountable participants ($M = 0.08$ vs. $M = 0.20$); this effect was not significant, however.

The one-factorial ANOVA testing the effect of type of accountability on the pattern of search was also not significant.

5. Concentration of search

The fact that the overall amount of information searched for each alternative was quite similar, already suggested by the low variability of search across alternatives, was also indicated by the concentration of search measures. Mean concentration of search across all conditions was 0.06 for the total number of items searched and 0.07 for different cells searched. The fact that these values are positive shows that the chosen alternative received more attention than the non-chosen alternatives. The difference to 0, the value for a completely balanced search, was significant both for concentration of search for the total number of items searched and for different cells searched (total number of items searched: $t(71) = 2.15$, $p = .035$, two-sided, different cells searched: $t(71) = 3.85$, $p = .000$)

In line with expectations, participants who had been made accountable tended to show a lower concentration of search for the total number of items searched than participants who had not been made accountable ($M = 0.01$ vs. $M = 0.11$, $t(70) = 1.81$, $p = .075$, two-sided). The difference for concentration of search for different cells searched was in the same direction, but not significant ($M = 0.05$ vs. $M = 0.09$).

The one-factorial ANOVAs employed to test the effect of type of accountability on the concentration of search measures suggested a very weak tendency of such an effect on the concentration of search for the total number of items searched ($F(2,69) = 1.89$, $p = .159$). Post-hoc simple contrasts showed that participants who had not been made accountable tended to concentrate their search more on the chosen alternative than participants who had been made internally accountable ($M = 0.11$ vs. $M = -0.02$, $p = .062$). No significant effects were obtained in the analysis performed on the concentration of search for different cells searched.

6. Search content

Given that one of the common effects of external accountability is to enforce salient norms, one may expect external accountability not only to have an effect on the decision process but also on the type of information searched. This may become evident as an increased attention to attributes which are considered to be particularly relevant for a justification in the given decision situation.

In order to test this assumption, one-factorial ANOVAs testing the effect of type of accountability on the percentage of items searched for the individually most important, second-most important, etc. attribute (relative to the total number of items searched) were performed. They revealed a very weak, non-significant effect of type of accountability on the search for the individually most important attribute ($F(2,69) = 1.66$, $p = .198$). Post-hoc simple contrasts showed that participants who had not been made accountable tended to search a relatively higher percentage of information for the attribute they considered most important than participants who had been made internally accountable ($M = 28\%$ vs. $M = 24\%$, $p = .09$, two-sided). The difference between the 'external accountability' condition ($M = 28\%$) and the other conditions was not significant.

In addition, in order to identify attributes that received different amounts of attention depending on the type of accountability participants experienced, one-factorial Analyses of Covariance (ANCOVAs) were performed on the percentage of items searched for each attribute (relative to the total number of items searched), holding the initial normalised weight for the particular attribute constant. These only revealed a marginally significant effect of type of accountability on the attribute ‘educational stay in a foreign country’ ($F(2,68) = 2.83$, $p = .066$). Post-hoc simple contrasts showed a significant difference between the ‘no accountability’ condition and the 'internal accountability' condition. Participants who had been made internally accountable searched a significantly higher percentage of items for ‘educational stay in a foreign country’ than participants who had not been made accountable (adjusted means: $M = 23\%$ vs. $M = 18\%$, $p = .021$). The difference to the 'external accountability' condition was not significant (adjusted means: $M = 23\%$ vs. $M = 19\%$). The covariate had a significant effect ($F(1,68) = 5.66$, $p = .020$).

Summary of the results for the first information search

Overall, the analysis of the processes observed when participants were making their first decision did not consistently support the predictions of the PMA. Although the effects of type of accountability on the various information search measures typically showed the expected pattern, namely a tendency for a deeper and less variable information search that was not concentrated on the eventually chosen alternative for participants who had been made accountable, compared to participants who had not been made accountable, these effects often failed to reach significance. This is likely to have been the consequence of a ceiling effect, induced by the demands of the task. Information search was generally exhaustive and therefore not very variable. Therefore, accountability could not lead to a more exhaustive search over and above the already exhaustive search induced by the task itself, especially not where the search of different cells was concerned. The ceiling effect could be the result of telling participants not only to choose one candidate but also to judge each candidate with respect to suitability for the job. Process tracing studies have shown that the required response mode can strongly influence information processing and that having to judge each alternative, compared to choosing one alternative, results in a deeper and more compensatory search (Billings & Scherer, 1988; Westenberg & Koele, 1992).

It is interesting to note that, although in the 'external accountability' condition, depth of search with respect to the total number of items searched was lowest, this group also showed the least variable search across alternatives and across attributes, where the total number of items searched was concerned. Trying to prepare for a justification and discussion with other people may have resulted in a better memory for information, making repeated searches less necessary and thereby keeping the total number of information items searched comparatively low. This assumption is supported by the fact that the mean percentage of repeatedly searched information was lowest in the external accountability group (M = 38%, no accountability: M = 43%, internal accountability: M = 43%), although the difference to the other conditions was not significant.

Pre-decisional external accountability, compared to no and internal accountability, also did not significantly enhance the attention to attributes that were considered most

important in the given decision situation. Instead, it was no accountability which resulted in the highest observed percentage of information searched for the most important attribute.

Information evaluation and integration

The following indicators of how participants evaluated and integrated the information presented to them were analysed:

- Choices

  Apart from testing whether the chosen alternative differed as a function of type of accountability, it was analysed whether there were any differences in the likelihood of a lexicographic choice between conditions, that is, a choice of the alternative with the best outcome on the most important attribute.
- Ratings of alternatives

  In addition to participants' suitability ratings for each alternative, the difference between their suitability rating for their chosen alternative and their mean rating for their non-chosen alternatives was analysed. This served as an indicator of the amount of differentiation at the first decision.
- Actual vs. predicted ratings of alternatives

  For each participant, the difference between the actual suitability rating for their chosen alternative and the MAUT predicted suitability rating for their chosen alternative was calculated. The same was done for the average non-chosen alternative. The latter value was then subtracted from the former, in order to arrive at a measure of the extent to which participants inflated the holistic evaluation of their chosen alternative more strongly than the holistic evaluation of their average non-chosen alternatives, indicating the extent to which they differentiated these alternatives from each other. The expected value of this measure is 0 if participants inflate the holistic evaluation of their chosen and average non-chosen alternative to the same extent, is > 0 if participants inflate the holistic evaluation of their chosen alternative relatively more than the holistic evaluation of their average non-chosen alternative, and is < 0 if they increase the holistic evaluation of their average non-chosen alternative relatively more than the holistic evaluation of their chosen alternative.

  The additive linear MAUT model specified in Chapter 1, Equation 6.1 was adapted to

allow for the case that participants did not search all available information in the following way:

$$E(A_i) = \frac{\sum_j S(x_{ij}) \cdot w_j \cdot v(x_{ij})}{\sum_j S(x_{ij}) \cdot w_j} \qquad \text{(Eq. 6.5)}$$

$S(x_{ij})$: search of the outcome of alternative i on attribute j, 0 = yes, 1 = no
w_j: weight for attribute j
$v(x_{ij})$: unidimensional value of the outcome of alternative i on attribute j

1. Choices

Table 6.4 presents the frequency of choice for each alternative, depending on type of accountability.

Table 6.4. Choice frequencies at the first decision (percentages in brackets).

	Type of Accountability			
Alternative	No	Internal	External	Σ
A	16 (44)	10 (55)	8 (44)	34 (47)
B	4 (11)	3 (17)	1 (6)	8 (11)
C	5 (14)	3 (17)	2 (11)	10 (14)
D	11 (31)	2 (11)	7 (39)	20 (28)
Σ	36 (100)	18 (100)	18 (100)	72 (100)

A Chi-squared analysis performed on the *overall* choice frequencies revealed that some alternatives were chosen significantly more often than others (χ^2 (3) = 23.56, p = .000). In particular, candidates A and D were clearly preferred to candidates B and C. Candidate A was chosen by 47% of the participants (34/72), candidate B by 11% of the participants (8/72), candidate C by 14% of the participants (10/72) and candidate D by 28% of the participants (20/72). Since the alternatives had been constructed in such a way that their overall attractiveness should have been similar if participants had applied a compensatory decision

rule, a non-compensatory decision rule seems to have been used by many participants. Consistent with this assumption, the two most popular candidates A and D were the alternatives with the best outcomes on the two attributes generally considered most important (number of placements in the area of Organisational Psychology, mean normalised weight $M = 0.28$, and job as a student research assistant, mean normalised weight $M = 0.24$).

It had been expected that participants who had been made externally accountable would be more likely to choose the alternative with the best outcome on what they considered to be the most important attribute, compared to internally accountable participants or participants who had not been made accountable. The frequency of lexicographic choices in each 'type of accountability' condition is shown in Table 6.5. As can be seen from this table, participants in the 'external accountability' condition were no more likely to make a lexicographic choice than participants in the other two conditions. Indeed, both a Chi-squared analysis comparing the three experimental conditions to each other and a Chi-squared analysis testing the 'external accountability' condition against a pooled condition of 'no accountability' and 'internal accountability' did not show any significant effects. If anything, the frequency of lexicographic choice was higher in the 'internal' compared to the 'external accountability' condition ($M = 67\%$ vs. $M = 56\%$).

Table 6.5. Lexicographic choice frequencies at the first decision (percentages in brackets).

Lexicographic Choice	Type of Accountability			
	No	Internal	External	Σ
No	15 (42)	6 (33)	8 (44)	29 (40)
Yes	21 (58)	12 (67)	10 (56)	43 (60)
Σ	36 (100)	18 (100)	18 (100)	72 (100)

2. Ratings of alternatives

The observed mean ratings of the suitability of each alternative reflected participants' choices. The highest overall mean rating was observed for candidate A ($M = 78.72$), followed by candidate D ($M = 72.17$), candidate C ($M = 61.28$) and candidate B ($M = 60.97$).

One-factorial ANOVAs testing the effect of type of accountability that were performed on the mean suitability ratings for each alternative, showed a significant effect on the rating of candidate D ($F(2,67) = 4.88$, $p = .010$). A simple post-hoc Bonferroni contrast indicated a significant difference between the 'no accountability' and the 'internal accountability' condition. Participants in the 'no accountability' condition rated candidate D as significantly more suitable than participants in the 'internal accountability' condition ($M = 76.94$ vs. $M = 60.28$, $p = .003$). This effect may partly be explained by different initial preferences. When the MAUT predicted overall evaluation of candidate D (based on the unidimensional values and weights elicited before the experimental manipulations) was introduced as a covariate, the effect of type of accountability was diminished ($F(2,68) = 2.87$, $p = .064$), but still marginally significant.

The overall mean rating for the chosen alternative was 89.44, the overall mean rating for the average non-chosen alternative was 62.62. As expected, the mean difference between these two ratings ($M = 13.69$) was highly significant ($t(71) = 16.33$, $p = .000$).

It had been expected that externally accountable participants would show a larger judgmental difference between their chosen and average non-chosen alternative than both internally accountable participants and participants who had not been made accountable. Figure 6.1 shows the obtained means in each condition. A one-factorial ANOVA testing the effect of type of accountability did not reveal any significant effect. Rather than being the highest, the mean obtained for the 'external accountability' condition was the lowest, suggesting that external accountability tended to result in a decreased rather than increased differentiation.

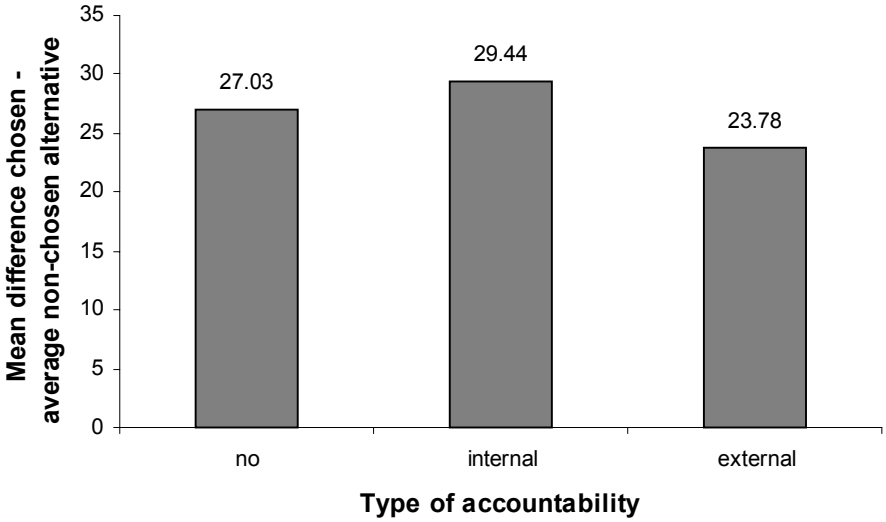


Figure 6.1. Mean difference in overall suitability between the chosen and the average non-chosen alternative after the first decision, as a function of type of accountability.

3. Actual vs. predicted ratings of alternatives

The overall difference between participants' actual and predicted rating for their chosen alternative and the actual and predicted rating for their average non-chosen alternative, which was calculated using the adapted MAUT model specified in Equation 6.5 above, was 13.56. This positive difference suggests that, overall, participants inflated the actual rating of the overall quality of their chosen alternative more strongly than their rating of the overall quality of their average non-chosen alternative.

It had been expected that for participants in the 'external accountability' condition, the obtained mean difference would be higher than for participants in the 'internal accountability' condition and the 'no accountability' condition. The results are presented in Figure 6.2. Contrary to expectations, the 'type of accountability' effect in the one-factorial ANOVA testing this assumption did not reach standard significance levels ($F(2,69) = 2.07$, $p = .134$). Also, the difference obtained in the 'external accountability' condition was the lowest. A planned contrast testing the 'external accountability' condition against the pooled other two conditions was marginally significant ($t(69) = 1.69$, $p = .096$, two-sided). This lends further support to the idea that participants in the 'external accountability' condition kept the

evaluative difference between their chosen alternative and their average non-chosen alternative smaller than participants in the 'internal accountability' condition and the 'no accountability' condition.

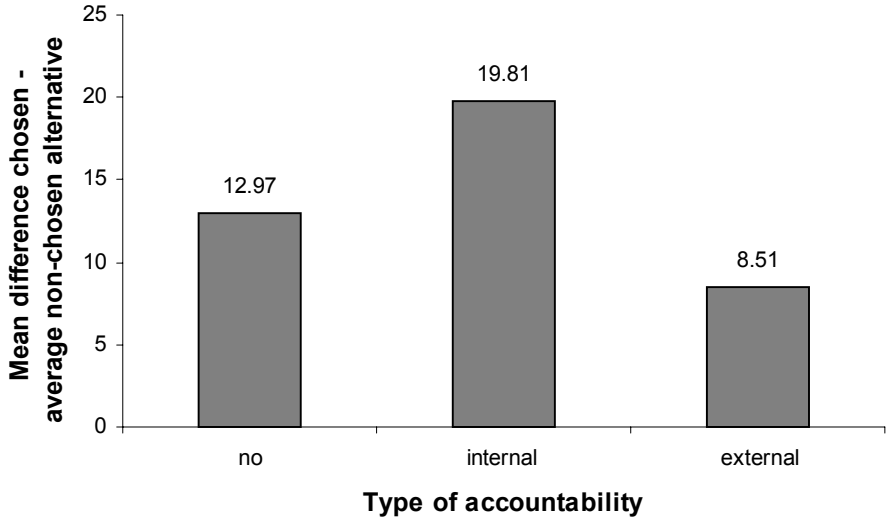


Figure 6.2. Mean difference between the actually observed and the MAUT predicted rating of overall suitability for the chosen vs. the average non-chosen alternative after the first decision, as a function of type of accountability.

Summary of the results for information evaluation and integration at the first decision

The expected effects of different types of accountability on the degree of bias in information and integration were not observed. It had been expected that externally accountable participants would show a stronger differentiation of their chosen alternative from their average non-chosen alternative and would be more likely make a lexicographic choice than internally accountable participants or participants who had not been made accountable. However, it was *internally* accountable participants who chose the alternative with the best outcome on the most important attribute most often, and evaluated information in such a way that the difference between their chosen alternative and average non-chosen alternative was greatest, while the tendency for externally accountable participants to show this behaviour was smallest. Even though the effects of type of accountability were not

significant for any of these measures, the consistency in the data suggests that it was indeed internal accountability that resulted in the relatively strongest bias. A more detailed discussion of the possible reasons for why this may have been the case will follow at the end of this chapter.

Final decision

Information search

The same measures as for the first information search were analysed. The results are shown in Table 6.6. In addition, since the attributes contained in information set 2 were largely redundant to those in information set 1, it was possible to test whether information search was selective to support the chosen alternative and, thereby, to consolidate the existing preference. If information search was biased in this way, it would be expected that for their chosen alternative, participants would search more information items for which they could anticipate a positive outcome than items for which they could anticipate a negative outcome. For their non-chosen alternatives, the opposite should be true. This selective information search should occur especially for information about important attributes, since important attributes would allow to discriminate between alternatives most effectively. A measure for biased search was developed that takes into account both the expected value of searched information and its importance, by comparing the mean product of unidimensional values and their corresponding weights for the information items searched for a particular alternative to the mean product that would be expected if information search was completely balanced. The biased search index (BSI) for an alternative A_i is defined as

$$BSI(A_i) = \frac{\sum_j FS(x_{ij}) \cdot w_{pj} \cdot v_p(x_{ij}) / \sum_j FS(x_{ij})}{(\sum_j w_{pj} \cdot v_p(x_{ij})) / m} \qquad \text{(Eq. 6.6)}$$

FS(xij): frequency of search of the outcome of alternative i on attribute j, i = 1, n; j = 1, m

wpj: weight for the parallel attribute of attribute j

vp(xij):	unidimensional value of the outcome of alternative i on the parallel attribute of attribute j
Σ wpj · v(xij):	MAUT prediction for the overall evaluation of alternative i, based on the attributes of information set1
m:	number of attributes (=5)

For an unbiased information search, a value of 1 is expected. A value > 1 indicates that information search was biased to support the searched alternative and a value < 1 indicates that it was biased to devalue the searched alternative. An overall measure of confirmation bias for the chosen alternative compared to either the competing alternative (the other stalemate alternative) or the non-chosen alternatives can be computed by dividing the BSI for the chosen alternative by the BSI for the competing alternative or the BSI for the average non-chosen alternative. A value > 1 is expected if there is a bias to support the chosen alternative over the competing or average non-chosen alternative, a value of < 1 if there is a bias to support the competing or average non-chosen alternative over the chosen alternative.

1. Depth of search

Information search was again quite extensive, but significantly less so than the first search (total number of information items searched: M = 109% vs. M = 191%, $t(71) = 6.96$, p = .000, different cells searched: M = 75% vs. M = 91%, $t(71) = 5.83$, p = .000). This may be explained by a tendency to concentrate the final information search mainly on the two alternatives between which a stalemate was said to have arisen.

In order to test this assumption, the percentage of information searched for these two alternatives was calculated, both for the total number of item searched and different cells searched. The overall means obtained were 69% and 68%, respectively. The fact that these values differed significantly from 50%, the expected value for a balanced search of all alternatives (total number of items searched: $t(71) = 7.55$, p = .000, different cells searched: $t(71) = 7.04$, p = .000), supported the hypothesis that information search was concentrated on the two stalemate alternatives.

Table 6.6. Process measures for the final information search.

	Time of Accountability						
	Before First Decision		Before Final Decision		*Significance of Effects*		
Type of Accountability	Internal	External	Internal	External	*Type*	*Time*	*Type x Time*
1. Depth of search (in %)							
total # of items	141	95	89	112			(*)
different cells	83	73	67	78			(*)
2. Variability of search across alternatives							
total # of items	1.75	1.53	2.10	2.15		(*)	
different cells	0.59	0.96	1.47	0.96		(*)	(*)
across attributes							
total # of items	0.95	1.31	0.89	1.02			
different cells	0.28	0.64	0.23	0.29	(*)	(*)	
3. Compensatory processing	0.70	0.51	0.37	0.56			(*)
4. Pattern of search (Payne Index)	-0.41	-0.45	-0.32	-0.21			
5. Concentration of search							
total # of items	0.16	0.32	0.65	0.37		(*)	(*)
different cells	0.18	0.34	0.60	0.33		*	*
6. Confirmation bias (BSI)							
chosen alternative	0.94	1.06	0.98	1.08	*		
competing alternative	1.06	0.97	0.93	1.01			(*)
aver. non-chosen alt.	1.01	0.94	0.98	1.05			*
chosen/competing alt.	0.92	1.12	1.03	1.06	*		
chosen/av. non-chosen	1.10	1.37	1.20	1.21			

Note: (*) $p \leq .10$
* $p \leq .05$

It had been expected that accountability before the final decision would result in a deeper final search than accountability before the first decision, yielding a main effect of time of accountability. In order to test the effects of type of accountability and time of accountability on depth of search, 2 (internal vs. external accountability) by 2 (accountability before the first decision vs. accountability before the final decision) ANOVAs were performed. The expected main effect of time of accountability could neither be observed for the percentage of total number of items searched nor for the percentage of different cells

searched. However, a marginally significant interaction was found between type of accountability and time of accountability for both measures (total number of items searched: $F(1,68) = 3.88$, $p = .053$, different cells searched: $F(1,68) = 3.32$, $p = .073$). Simple main effects analyses suggested that there was only a significant 'time of accountability' effect for participants who had been made internally accountable (total number of items searched: $F(1,68) = 4.42$, $p = .039$, different cells searched: $F(1,68) = 3.77$, $p = .056$). This effect was not in line with expectations, however. Internal accountability resulted in a significantly higher percentage of information searched, when it had already been introduced before the *first* decision than when it had been introduced before the final decision (total number of items searched: $M = 141\%$ vs. $M = 89\%$, different cells searched: $M = 83\%$ vs. $M = 67\%$). For external accountability, the opposite trend could be observed, but was not significant (total number of items searched: $M = 95\%$ vs. $M = 112\%$, different cells searched: $M = 73\%$ vs. $M = 78\%$).

Figure 6.3 shows the percentage of total number of items searched, as a function of type and time of accountability.

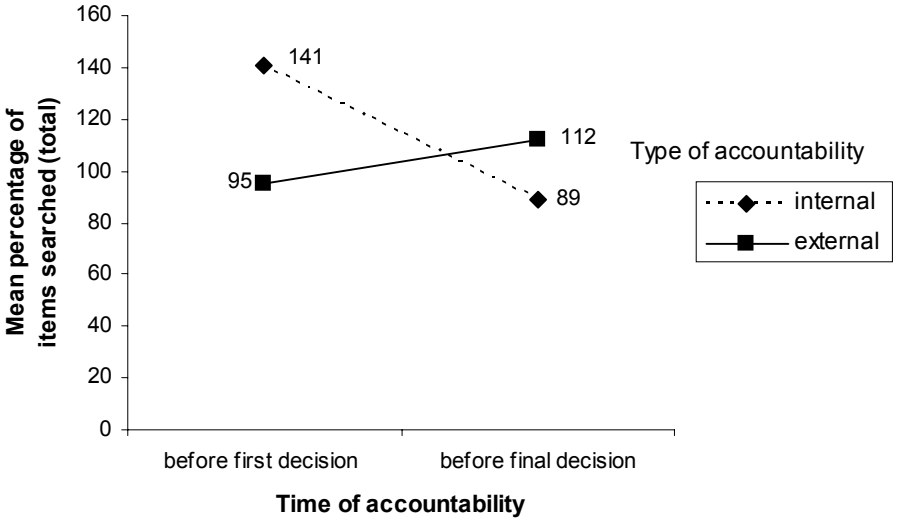


Figure 6.3. Mean percentage of items searched during the final information search (based on the total number of items searched), as a function of type of accountability and time of accountability.

Two-factorial ANOVAs were also performed on the percentage of information searched for the two stalemate alternatives only. The analyses yielded a marginally significant main effect of time of accountability for both the total number of items searched and different cells searched (total number of items searched: $F(1,68) = 3.78$, $p = .056$, different cells searched: $F(1,68) = 2.74$, $p = .100$). As expected, participants with accountability introduced before the final decision tended to pay more attention to the two stalemate alternatives than participants with accountability introduced before the first decision (total number of items searched: $M = 74\%$ vs. $M = 64\%$, different cells searched: $M = 72\%$ vs. $M = 64\%$). Simple main effects analyses suggested, however, that both for the total number of items searched and different cells searched, the significant effect of time of accountability was confined to the 'internal accountability' conditions (total number of items searched: $F(1,68) = 5.56$, $p = .021$, different cells searched: $F(1,68) = 5.87$, $p = .018$). Participants who had been made internally accountable before the final decision searched a higher percentage of information for the stalemate alternatives than participants who had been made internally accountable before the first decision (total number of items searched: $M = 78\%$ vs. $M = 61\%$, different cells searched: $M = 78\%$ vs. $M = 61\%$). For externally accountable participants, there was no significant difference (total number of items searched: $M = 70\%$ vs. $M = 67\%$, different cells searched: $M = 67\%$ vs. $M = 67\%$). This pattern of data suggests that under external accountability, more alternatives were considered, no matter whether accountability had been introduced before the first or before the final decision.

2. Variability of search

The overall mean variability of search across alternatives did not change significantly from the first to the final search where variability across alternatives for the total number of items searched was concerned ($M = 1.94$ vs. $M = 1.88$, $t(71) = 0.30$, *ns*), but was significantly larger during the final information search for different cells searched ($M = 0.32$ vs. $M = 1.00$, $t(71) = 5.22$, $p = .000$). For the overall mean variability of search across attributes, the opposite was true; there was no significant difference between the first and the final search for different cells searched ($M = 0.39$ vs. $M = 0.36$, $t(71) = 5.22$, *ns*), but variability of search across attributes calculated for the total number of items searched was significantly smaller for the final compared to the first information search ($M = 2.52$ vs. $M = 1.04$, $t(71) = 5.97$, p

= .000). Unlike for the first search, for the final search variability of search across alternatives was higher than variability across attributes (total number of items searched: M = 1.88 vs. M = 1.04, $t(71) = 5.41$, p = .000, different cells searched: M = 1.00 vs. M = 0.36, $t(71) = 4.85$, p = .000). This pattern of results is again consistent with a search behaviour that concentrates on only a subset of the available alternatives, for which information is searched on a similar number of attributes.

It had been expected that variability of search across alternatives and attributes would be smaller for participants who had been made accountable before the final compared to the first decision. Two-factorial ANOVAs testing the effects of type of accountability and time of accountability on the variability of search measures revealed a marginally significant main effect of time of accountability for variability of search across alternatives. However, contrary to predictions, participants who had been made accountable before the final decision showed a higher variability of search across alternatives than participants who had been made accountable before the first decision (total number of items searched: M = 2.12 vs. M = 1.64, $F(1,68) = 3.69$, p = .059), different cells searched: M = 1.22 vs. M = 0.77, $F(1,68) = 3.72$, p = .058). For variability across alternatives for different cells searched, this main effect was qualified by a marginally significant interaction between type of accountability and time of accountability ($F(1,68) = 3.71$, p = .058). Simple main effects analyses indicated that the 'time of accountability' effect was only significant for internally accountable participants ($F(1,68)$ = 7.43, p = .008). They showed a significantly lower variability of search across alternatives when accountability had already been introduced before the first decision, compared to when accountability had been introduced before the final decision (M = 0.59 vs. M = 1.47). For externally accountable participants, there was no difference (M = 0.96 vs. M = 0.96). This was consistent with the results observed for depth of search, but, again, contradicted the original predictions. This pattern of results would be expected, however, if participants who had been made accountable immediately prior to the final information search mainly focused on the two stalemate alternatives, as this would have created a high variability of search across alternatives which, in turn, would have prevented the expected main effect of time of accountability from occurring.

For variability of search across attributes for different cells searched, marginally significant main effects of type of accountability ($F(1,68) = 3.40$, p = .070) and of time of

accountability ($F(1,68) = 3.14$, $p = .081$) were observed. Participants who had been made internally accountable tended to show a lower variability of search across attributes than participants who had been made externally accountable ($M = 0.26$ vs. $M = 0.46$), and participants with accountability introduced before the first decision tended to show a higher variability of search across attributes than participants with accountability introduced before the final decision ($M = 0.46$ vs. $M = 0.26$). The latter result was consistent with predictions. Although the interaction between the two factors did not reach standard significance levels ($F(1,68) = 1.86$, $p = .178$), simple main effects analyses showed that there was a significant effect of time of accountability for participants who had been made externally accountable only ($F(1,68) = 4.91$, $p = .030$). These participants showed a lower variability of search when they had been made accountable before the final decision than when they had been made accountable before the first decision ($M = 0.29$ vs. $M = 0.64$). For participants in the 'internal accountability' conditions, the difference was not significant ($M = 0.23$ vs. $M = 0.28$). For variability of search across attributes for the total number of items searched, no significant effects were obtained.

The results for variability of search across attributes for different cells searched were consistent with predictions and supported one of the earlier conclusions, based on the results observed for depth of search for the two stalemate alternatives; external accountability in particular seemed to have had the effect of making participants consider more alternatives.

3. Compensatory processing

The overall mean value for the Index of Compensatory Processing suggested that, in general, information processing was less compensatory for the final than for the first search ($M = 0.82$ vs. $M = 0.54$, $t(71) = 5.67$, $p = .000$). This may again be attributed to the fact that participants' attention was focused on their chosen and the competing stalemate alternative.

Consistent with the results for depth of search and variability of search across alternatives for different cells searched, the two-factorial ANOVA carried out on the index of compensatory processing yielded a marginally significant interaction between type of accountability and time of accountability ($F(1,68) = 3.88$, $p = .053$). Simple main effects analyses indicated that the effect of time of accountability was only significant for internally

accountable participants ($F(1,68) = 6.05$, $p = .016$). Their information processing was significantly more compensatory when accountability had been introduced before the first decision than when it had been introduced before the final decision ($M = 0.70$ vs. $M = 0.37$). For external accountability, there was no significant difference ($M = 0.51$ vs. $M = 0.56$). Again, this pattern of results would be expected if participants who had been made internally accountable before the final decision concentrated their search on the two stalemate alternatives.

4. Pattern of search (Payne Index)

The Payne Index suggested that the final information search was generally more non-compensatory than the first search, given that it was significantly more attribute-wise than the first search ($M = -0.35$ vs. $M = 0.14$, $t(71) = 6.69$, $p = .000$). This time the difference to 0.05, the expected value for a balanced search, was significant ($t(71) = 5.82$, $p = .000$).

It had been expected that participants who had been made accountable before the final decision would show a less attribute-wise search than participants who had already been made accountable before the first decision. Although the observed means followed this direction ($M = -0.26$ vs. $M = -0.43$), the main effect of time of accountability in the two-factorial ANOVA carried out on the data was not significant. The main effect of type of accountability and the interaction between time of accountability and type of accountability were also not significant.

5. Concentration of search

The concentration of search on the chosen alternative was higher during the final than during the first search (total number of information items searched: $M = 0.38$ vs. $M = 0.06$, $t(71) = 5.22$, $p = .000$, different cells searched: $M = 0.36$ vs. $M = 0.07$, $t(71) = 5.22$, $p = .000$). This finding was once more consistent with the assumption that only a subset of alternatives that included the chosen alternative had been searched. This is, because the concentration of search measure relates the number of searched items that would be expected if all alternatives were searched to the same extent as the chosen alternative to the actual number of items searched. If only some of the available alternatives are searched, a high value is obtained.

It had been expected that concentration of search would be lower for participants who had been made accountable before the final decision than for participants who had been made accountable before the first decision, indicated by a significant main effect of time of accountability. For concentration of search for the total number of items searched, the predicted main effect of time of accountability was found ($F(1,68) = 6.04$, $p = .017$). However, unlike expected, participants with accountability introduced before the final decision showed a significantly *higher* concentration of search than participants with accountability introduced before the first decision ($M = 0.51$ vs. $M = 0.24$). No significant main effect emerged for the 'type of accountability' factor. There was, however, a marginally significant interaction between type of accountability and time of accountability ($F(1,68) = 3.84$, $p = .054$). Simple main effect analyses showed that the effect of time of accountability was only significant in the 'internal accountability' conditions ($F(1,68) = 9.76$, $p = .003$). Internally accountable participants showed a significantly higher concentration of search on the chosen alternative for the total number of items searched when they had been made accountable before the final compared to the first decision ($M = 0.65$ vs. $M = 0.16$), whereas for externally accountable participants this difference was not significant ($M = 0.37$ vs. $M = 0.32$).

The results for concentration of search for different cells mirrored those for the total number of items searched; there was a significant main effect for time of accountability ($F(1,68) = 4.62$, $p = .035$) such that participants with accountability introduced before the final decision showed a significantly higher concentration of search on the chosen alternative than participants with accountability introduced before the first decision ($M = 0.47$ vs. $M = 0.26$), and a significant interaction between type of accountability and time of accountability ($F(1,68) = 4.71$, $p = .033$). Simple main effect analyses showed that time of accountability only had a significant effect on internally accountable participants ($F(1,68) = 9.33$, $p = .003$). Concentration of search for different cells was significantly higher when internal accountability had been introduced before the final compared to when it had been introduced before the first decision ($M = 0.60$ vs. $M = 0.18$). There was no significant difference for externally accountable participants ($M = 0.33$ vs. $M = 0.34$). A marginally significant difference was found between internally and externally accountable participants when accountability had been introduced before the final decision ($M = 0.60$ vs. $M = 0.33$, $F(1,68)$

= 3.71, p = .058), but not when it had been introduced before the first decision (M = 0.18 vs. M = 0.34).

6. Search content

Like for the first search, it was checked whether the accountability manipulations had affected the attention to particular attributes. For this purpose, two-factorial ANOVAs were performed on the percentage of items searched for the individually most important, second-most important, etc., attribute in information set 2 (related to the total number of items searched). They revealed a marginally significant effect of type of accountability for the individually least important attribute (F(1,68) = 3.04, p = .086). Externally accountable participants showed a tendency to search a lower percentage of items for the attribute they considered least important than internally accountable participants (M = 16% vs. M = 19%). These results are consistent with the assumption that external accountability makes people focus more on what they assume to be the important aspects in a given decision situation, at the cost of less important aspects.

In order to identify attributes that had received different amounts of attention from participants in the different experimental conditions, two-factorial ANCOVAs, testing the effects of type of accountability and time of accountability, were performed on the percentage of information searched for each attribute, holding the normalised weight for the analysed attribute constant. These did not reveal any significant effects.

When the percentage of different cells instead of the total number of items searched was examined for each attribute, a marginally significant effect of type of accountability was found for the attribute ‘student representative on university committees’ (F(1,67) = 3.09, p = .084). Participants experiencing internal accountability tended to search a higher percentage of cells for this attribute than participants with external accountability (M = 20% vs. M = 17%). The covariate also had a significant effect (F(1,67) = 6.46, p = .013).

7. Confirmation bias

Contrary to expectations, the final information search was not generally biased to support the chosen alternative or devalue the competing or the average non-chosen alternative. One-sample t-tests, testing the observed overall means of the Biased Search Index (BSI) for the chosen alternative (M = 1.01, $t(71)$ = 0.57, *ns*), the competing alternative (M = 0.99, $t(71)$ = -0.39, *ns*) and the average non-chosen alternative (M = 1.00, $t(71)$ = -0.23, *ns*) against a value of 1, which would be expected if information search was unbiased, revealed no significant effects.

The mean ratio of the BSI for the chosen alternative to the BSI for the average non-chosen alternative (M = 1.22), however, was significantly different from 1 ($t(71)$ = 4.54, p = .000), suggesting that information search was biased to support the chosen alternative over the average non-chosen alternative. The mean ratio of the BSI for the chosen alternative to the BSI for the direct competitor (M = 1.03), on the other hand, was not significantly different from 1. This may be explained by the fact that the final information search seemed to have been strongly focused on the two stalemate alternatives. If those two alternatives had been searched equally, a ratio near 1 would have been obtained, indicating an unbiased search.

In order to test the effect of type of accountability and time of accountability on the tendency to exhibit a confirmation bias, two-factorial ANOVAs were performed on the biased search index (BSI) for the chosen alternative, the competing alternative and the average non-chosen alternative. It had been expected that the information search of participants who had been made accountable before the final decision would be more biased to support the chosen and to devalue the competing or average non-chosen alternative than that of participants who had been made accountable before the first decision. For the chosen alternative, the expected significant effect of time of accountability was not found, but the main effect of type of accountability was significant. For externally accountable participants, a significantly higher mean value was observed than for internally accountable participants (M = 1.07 vs. M = 0.96, $F(1,68)$ = 5.03, p = .028). One-sample t-tests testing the observed means against a value of 1 showed that the BSI for externally accountable participants was significantly greater than 1, indicating that their information search was biased to support the chosen alternative ($t(35)$ = 1.68, p = .050, one-sided). For internally accountable participants, a marginally significant

difference was observed, indicating a tendency of their information search to be biased to devalue the chosen alternative ($t(35) = -1.58$, p = .061, one-sided). These findings are illustrated in Figure 6.4.

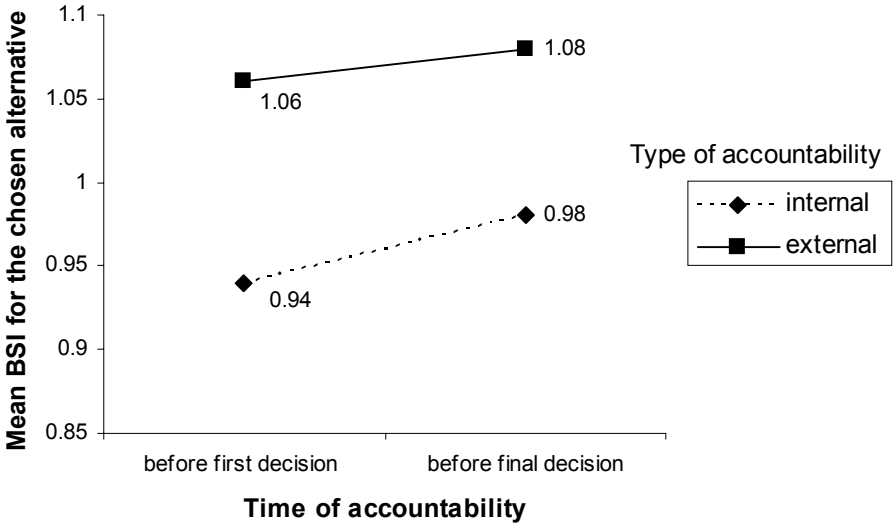


Figure 6.4. Mean Biased Search Index for the chosen alternative (a value >1 indicates a bias to support the chosen alternative, a value <1 indicates a bias to devalue the chosen alternative).

The ANOVA analysis for the competing alternative revealed a marginally significant interaction between time and type of accountability (F (1,68) = 3.63, p = .061). Simple main effects analyses suggested that the effect of time of accountability was only significant for participants in the 'internal accountability' conditions ($F(1,68) = 4.25$, p = .043). Internal accountability before the final decision resulted in a significantly stronger bias to devalue the competing alternative than internal accountability before the first decision (M = 0.93 vs. M = 1.06). This was consistent with expectations. For external accountability, the opposite pattern, although not significant, could be observed (M = 1.01 vs. M = 0.97). For the average non-chosen alternative there was also a significant interaction effect between time of accountability and type of accountability ($F(1,68) = 4.11$, p = .047). This time, however, simple main effects analyses suggested that the effect of time of accountability was only significant for *externally* accountable participants ($F(1,68) = 5.20$, p = .026). Participants who had been made externally accountable before the first decision showed a significantly stronger

bias to devalue the average non-chosen alternative than participants who had been made accountable before the first decision (M = 0.94 vs. M = 1.05, see Figure 6.5).

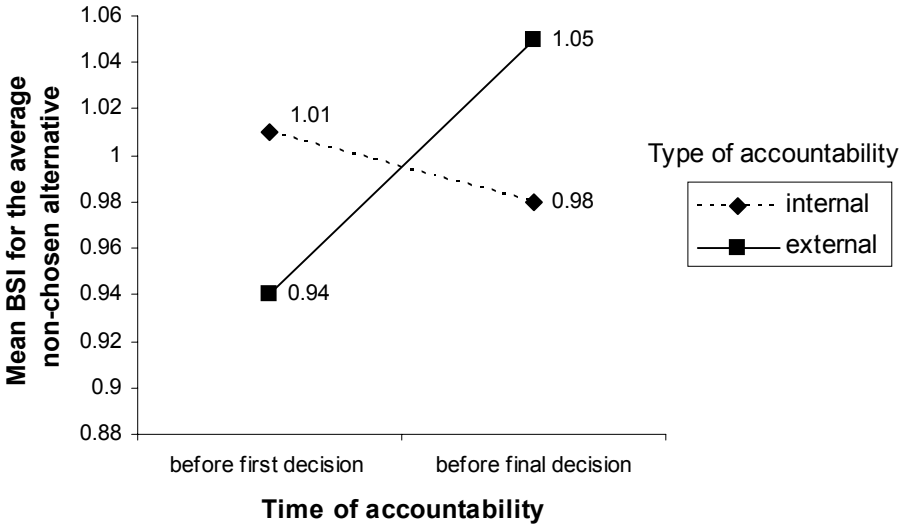


Figure 6.5. Mean Biased Search Index for the average non-chosen alternative (a value >1 indicates a bias to support the average non-chosen alternative, a value <1 indicates a bias to devalue the average non-chosen alternative).

The ANOVA performed on the ratio of the BSI for the chosen to the BSI for the competing alternative yielded a significant effect of type of accountability (F (1,68) = 4.68, p = .034). The information search of externally accountable participants showed a stronger overall bias to support the chosen compared to the competing alternative than the information search of internally accountable participants (M = 1.09 vs. M = 0.97). The analysis on the ratio of the BSI for the chosen to the BSI for the average non-chosen alternative did not yield any significant effects.

The above results were obtained for the whole sample of participants. It may be argued that the confirmation bias can be expected to be stronger in participants who did not change their initial choice when making their final decision. Analyses that only included data from participants who kept their previously chosen alternative supported this assumption. The overall pattern of results remained unchanged, however. The only difference observed was that the marginally significant interaction between type of accountability and time of accountability in the analysis of the BSI for the competing alternative became significant.

Summary of the results for the final information search

Generally, the results for the measures of information search observed during the final decision were not consistent with the predictions at the outset. In cases where the expected pattern was found, a large individual response variability meant that the effects often failed to reach statistical significance. It may also be argued, however, that particular features of the task meant that the hypothesised effects were not observed.

For example, it had been expected that the information search of participants who had been made accountable immediately prior to the final decision would be deeper, less variable and more compensatory than that of participants who had already been made accountable before the first decision. This effect was not generally observed, except in the variability of search across attributes for different cells searched, where participants who had been made accountable before the final decision showed the expected smaller variability of search across attributes, suggesting that their information processing was indeed more compensatory compared to that of participants who had been made accountable before the first decision. Other measures, however, for example, the variability of search across alternatives and the concentration of search on the chosen alternative, seemed to contradict this finding and seemed to indicate that the information search of participants who had only been made accountable before the final decision was *less* compensatory compared to that of participants who had been made accountable before the first decision. Also, significant interaction effects between time of accountability and type of accountability on these measures suggested that this was particularly the case for participants who had been made internally accountable before the final decision. The apparent contradiction in results can be resolved when one considers that, before their final information search, participants were told that a stalemate had arisen between two job candidates. One may therefore assume that participants particularly concentrated their search on these two candidates. This assumption was supported by the fact that the final information search was generally less extensive than the first and that about 70% of the searched information was for the two stalemate alternatives (69% for total number of items searched and 68% for different cells searched). Also, participants who had been made accountable before the final decision tended to focus more on the two stalemate alternatives than participants made accountable before the first decision, especially when they were internally accountable. This concentration on the stalemate alternatives would have

made information search less extensive and more variable, especially where variability of search across alternatives was concerned. The pattern of results for the final information search may therefore be explained by a tendency of participants who had been made accountable immediately prior to this search to concentrate on the two stalemate alternatives. This is consistent with the PMA model, in that accountability intensified information search for the alternatives that had been brought to the special attention of participants. A reason for the fact that this effect was stronger in the 'internal accountability' compared to the 'external accountability' condition may be that externally accountable participants anticipated a justification of their decision to their fellow committee members and therefore may have also looked for comparison information about other alternatives, in order to be well-prepared for a discussion.

Analyses of the content of the searched information indicated that externally accountable participants showed a tendency to search a lower percentage of items for the attribute they individually found least important than internally accountable participants. Also, externally accountable participants tended to search information about the ‘student representative on university committees’ attribute less frequently than participants with internal accountability. This attribute had received the lowest mean importance rating of all attributes in information set 2 ($M = 0.16$). Both findings are consistent with the prediction that external accountability will make external norms salient and will focus the individual's attention on what they perceive to be important in the situation, to the detriment of less important attributes. Also, the 'student representative' attribute may have been tempting for student participants to take into account, as it was the attribute that told most about a candidate’s interest in and commitment to students’ issues, but may not have been perceived as relevant to the professors and junior lecturers who were also said to be part of the selection committee. The PMA predicts that, as this external norm should be particularly salient to externally accountable participants, they should not consider this type of information as strongly as internally accountable participants. This is what was observed.

The PMA had also predicted that participants' final information search would show signs of a confirmation bias, in that participants would predominantly search information that they could expect to support their previous choice. While the Biased Search Index (BSI) that was calculated to measure this tendency did not suggest a significantly biased information

search for the chosen alternative overall, the tendency to search more information that supported rather than did not support the chosen alternative was significant in externally accountable participants' information search, and significantly stronger for externally than internally accountable participants. The expected main effect of time of accountability was not obtained. This was also true for the ratio of the BSI for the chosen to the competing alternative which indicates the extent to which participants searched more supporting information for their chosen compared to their competing alternative. The significant main effects of type of accountability may be explained by assuming that for externally accountable participants, the conflict between their own choice and that favoured by some other members of the committee was more salient, as they anticipated having to publicly justify their choice later on and therefore may have felt a need to bolster their chosen alternative more than internally accountable participants, while at the same time also looking for the advantages of the other alternatives. Indeed, the analysis of the BSI for the average non-chosen alternative suggested that participants who had been made externally accountable before the final decision showed a significantly less strong bias to search information that devalued the average non-chosen alternative than participants who had been made externally accountable before the first decision. Participants who had been made internally accountable before the final decision were found to search significantly more information that devalued the competing alternative compared to participants who had been made internally accountable already before the first decision.

Information evaluation and integration

The following main types of dependent variables were analysed:

- Choices

  Changes of preferences as a function of time of accountability and type of accountability were analysed.

- Attribute weights

  For each participant, the difference between the normalised weights obtained for each attribute in information set 1 after the final and before the first decision was calculated, in order to test whether importance differentiation had occurred. For this to be the case, the weights of important attributes on which the chosen alternative was likely to have a

positive outcome would have to be increased, and the weights of less important attributes on which the chosen alternative was likely to have a negative outcome would have to be decreased, resulting in a positive difference for important attributes and a negative difference for less important attributes.

- Ratings of alternatives

  Apart from the difference between participants' suitability rating for their chosen and average non-chosen alternative, the difference between participants' suitability ratings for their chosen alternative and the competing alternative (the other stalemate alternative) was analysed. Also, the change in these measures from the first to the second elicitation was analysed.

- MAUT predictions

  MAUT predictions of the overall evaluation of alternatives were used to test further hypotheses regarding differentiation and consolidation processes. Since MAUT predictions take into account *both* the importance of attributes and the evaluation of attribute outcomes, they allow to assess the joint effects of importance differentiation and attractiveness differentiation. It is assumed that both differentiation processes are synchronised, because, in order to enhance the overall evaluation of one's chosen alternative compared to the competing alternative or, more generally, the average non-chosen alternative, it is most effective to change one's perception of the attractiveness of outcomes for attributes with a high compared to a low importance. MAUT predictions were calculated using the adapted additive linear model specified in Equation 6.5. Only the attributes of information set 1 were considered, since only for them unidimensional values and weights had been elicited before the first and after the final decision.

  The dependent variables analysed were the difference between the MAUT predicted overall evaluation of the chosen and the average non-chosen alternative and the difference between the MAUT predicted overall evaluation of the chosen and the competing alternative as well as the change in these two difference values from the first to the second elicitation. The first two values are indicators of the amount of consolidation that has taken place, the latter two represent the sum of both differentiation and consolidation effects.

1. Choices

Table 6.7 presents the choice frequencies for each alternative at the first and final decision. If participants' second choice had been the same as their first choice, all values except in the diagonal should have been 0. This, however, was not the case. Candidate A was chosen less frequently than at the first decision (36% instead of 47%), whereas the choice frequencies for the other candidates increased slightly, especially for candidate D (B: 12% vs. 11%, C: 17% vs. 14%, D: 35% vs. 28%).

Although, in order to allow participants to keep their choice, the stalemate had been said to involve each participant's preferred alternative, a considerable number of participants (14/72 = 19%) changed their previously chosen alternative (see Table 6.8). Thirty-four participants (47%) changed their second-best alternative, 13 of those (18%) changed both their chosen and their second-best alternative. Eleven participants (15%) made the competing alternative their first choice, whereas 15 participants (21%) made it their second-best alternative.

Table 6.7. Choice frequencies for each alternative at the first and final decision (percentages in brackets).

	Chosen Alternative, Final Decision				
Chosen Alternative, First Decision	A	B	C	D	Σ
A	24	3	1	6	34 (47)
B	-	6	1	1	8 (11)
C	-	-	10	-	10 (14)
D	2	-	-	18	20 (28)
Σ	26 (36)	9 (12)	12 (17)	25 (35)	72 (100)

There were neither any significant differences between conditions in terms of whether participants had changed their preferred alternative or not (see Table 6.8), nor which alternative had been preferred at the final decision.

Table 6.8. Frequency of change of the chosen alternative, as a function of type of accountability and time of accountability.

	Time of Accountability				
	Before First Decision		Before Final Decision		
Type of Acc. / Change Of Choice	Internal	External	Internal	External	Σ
No	16	14	14	14	58 (81)
Yes	2	4	4	4	14 (19)
Σ	18	18	18	18	72 (100)

2. Attribute weights

Table 6.9 presents the mean normalised weights for the individually most important, second-most important, etc. attribute (based on the first importance ratings) after the final decision, before the first decision, and the difference between the two. Since importance differentiation and consolidation would have been difficult for participants who had changed their chosen alternative, because alternatives had been constructed in such a way that each alternative had their best outcome on a different attribute, only the data of those participants who had not done so were included in the following analyses.

Overall, the results did not point to importance differentiation and consolidation having taken place. Instead, the difference between the first and second weight elicitation indicated a mere regression effect; the more important weights were decreased, resulting in negative differences between the second and first elicitation, and the less important weights were increased, resulting in positive differences between the second and first elicitation.

Table 6.9. Mean normalised weights before the first and after the final decision, as a function of type of accountability and time of accountability.

	Time of Accountability						
	Before First Decision		Before Final Decision		*Significance of Effects*		
Type of Accountability Rank of Importance	Internal (n=16)	External (n=14)	Internal (n=14)	External (n=14)	*Type*	*Time*	*Type x Time*
1							
after final decision	0.295	0.276	0.281	0.302			
before first decision	0.322	0.326	0.291	0.337			
difference	-0.027	-0.050	-0.010	-0.035	(*)		
2							
after final decision	0.244	0.248	0.251	0.259			
before first decision	0.265	0.264	0.255	0.278			
difference	-0.021	-0.017	-0.004	-0.018			
3							
after final decision	0.207	0.232	0.211	0.193			(*)
before first decision	0.217	0.197	0.219	0.196	*		
difference	-0.010	0.036	-0.008	-0.003	*		(*)
4							
after final decision	0.167	0.153	0.171	0.154			
before first decision	0.136	0.147	0.167	0.131			
difference	0.031	0.006	0.004	0.024			
5							
after final decision	0.087	0.091	0.087	0.091			
before first decision	0.060	0.066	0.068	0.063			
difference	0.027	0.026	0.018	0.029			

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.
(*) $p \leq .10$
* $p \leq .05$

Two-factorial ANOVAs testing the effects of type of accountability and time of accountability on the *difference* values revealed a marginally significant main effect of type of accountability for the individually most important attribute ($F(1,54) = 3.65$, $p = .061$). Participants under internal accountability tended to reduce their most important weight to a lesser extent than participants under external accountability ($M = -0.012$ vs. $M = -0.043$). This effect was largely reduced, however, and not significant anymore, if the weight for the most

important attribute at the first elicitation was introduced as a covariate, thereby holding any differences between conditions before the experimental manipulations constant ($F(1,53) = 1.77$, $p = .189$). The covariate was highly significant ($F(1,53) = 9.50$, $p = .003$). The expected main effect of time of accountability was not obtained.

In addition, a significant main effect of type of accountability was obtained in the analysis of the individually third-most important attribute ($F(1,54) = 5.41$, $p = .024$). Whereas participants in the 'internal accountability' conditions reduced the weight for the third-most important attribute slightly from the first to the second elicitation, participants in the 'external accountability' conditions increased it, indicating importance differentiation and consolidation ($M = -0.009$ vs. $M = 0.017$). Although a significant difference between internally accountable and externally accountable participants existed before the first decision (the third-most important weight was significantly higher for internally accountable participants than for externally accountable participants, $M = 0.22$ vs. $M = 0.20$), when the value obtained at the first elicitation was introduced as a covariate in an ANCOVA, the covariate effect of the initial third-most important weight was not significant. The main effect of type of accountability was reduced, but remained marginally significant ($F(1,53) = 3.21$, $p = .079$). In addition to the main effect of type of accountability, there was a marginally significant interaction between type of accountability and time of accountability ($F(1,54) = 3.34$, $p = .073$). This remained almost unchanged in the ANCOVA ($F(1,53) = 3.21$, $p = .071$). Simple main effects analyses showed that there was only a significant effect of time of accountability when participants were externally accountable ($F(1,54) = 5.73$, $p = .020$). For participants who had been made externally accountable before the final decision the difference in their third-most important weight between the second and the first elicitation was significantly smaller than for participants who had been made externally accountable before the first decision ($M = -0.003$ vs. $M = 0.036$). This was opposite to what had been expected. Also, the effect of type of accountability was only significant for participants who had been made accountable before the first decision ($F(1,54) = 8.91$, $p = .004$). Participants who had been made externally accountable before the first decision showed a larger difference than participants who had been made internally accountable before the first decision ($M = 0.036$ vs. $M = -0.01$).

Two-factorial ANOVAs employed to test the effects of type of accountability and time of accountability on the differences between the second and first weight judgement for the specific attributes presented in information set 1 (rather than individually ranked weights) revealed no significant effects.

3. Overall evaluation of alternatives

Table 6.10 presents the results for the evaluation measures taken to assess the joint effects of importance and attractiveness differentiation. As before, only the data of participants who had not changed their choice were included in the analysis.

3a. Ratings of the chosen vs. the average non-chosen alternative

The results showed that, with the exception of participants who had been made externally accountable before the final decision, participants did not show a tendency to consolidate the judgmental difference between their chosen alternative and their average non-chosen alternative, but decreased this difference slightly instead. The overall mean difference after the first decision was 27.86, and 26.05 after the final decision, resulting in a negative overall mean difference of -1.81 between the first decision and the final decision. This value was not significantly different from 0.

Two (type of accountability: internal vs. external) by two (time of accountability: before the first vs. before the final decision) ANOVAs performed on the judgmental difference between the chosen alternative and the average non-chosen alternative at the first decision and at the final decision did not reveal any significant effects.

The ANOVA performed on the difference between the first and the second elicitation revealed a weak tendency for a 'time of accountability' main effect ($F(1,54) = 2.49$, $p = .121$). Participants who had been made accountable before the final decision tended to decrease the difference between their chosen alternative and their average non-chosen alternative to a smaller extent than participants who had been made accountable before the first decision ($M = 0.38$ vs. $M = -3.86$). No other effect was significant.

Table 6.10. Measures of the evaluative difference between alternatives after the final decision.

	Time of Accountability						
	Before First Decision		Before Final Decision		*Significance of Effects*		
Type of Accountability	Internal (n=16)	External (n=14)	Internal (n=14)	External (n=14)	*Type*	*Time*	*Type x Time*
1. Rating chosen - av. non-chosen alternative							
after final decision	25.94	22.38	25.21	30.67			
after first decision	31.46	24.33	25.48	29.64			
difference	-5.52	-1.95	-0.26	1.02			
2. MAUT chosen - av. non-chosen alternative							
after final decision	10.26	10.40	15.84	17.46		(*)	
before first decision	11.54	11.08	12.17	11.33			
difference	-1.28	-0.69	3.66	6.13			
3. Rating chosen - competing alternative							
after final decision	21.56	15.00	22.57	23.14			
after first decision	34.06	23.43	29.79	34.64			
difference	-12.50	-8.43	-7.21	-11.50			
4. MAUT chosen - competing alternative							
after final decision	8.43	11.23	16.20	17.26			
before first decision	11.87	7.99	13.54	11.72			
difference	-3.44	3.24	2.66	5.53			

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.
(*) $p \leq .10$
* $p \leq .05$

3b. MAUT predictions for the chosen vs. average non-chosen alternative

The overall mean difference in MAUT predictions between participants' chosen alternative and their average non-chosen alternative indicated that a small degree of differentiation and consolidation had occurred. The overall mean difference before the first decision was 11.53, after the final decision it was 13.38. The obtained overall mean difference of 1.85 between the difference in MAUT predictions before the first and after the final decision was not significantly different from 0, however.

The two-factorial ANOVA performed on the difference in MAUT predictions for the chosen alternative compared to the average non-chosen alternative at the *first elicitation* (before the experimental manipulations), as expected, revealed no significant differences between conditions. For the difference at the *second elicitation* after the final decision, however, a marginally significant main effect of time of accountability ($F(1,54) = 3.86$, $p = .055$) was found. Consistent with expectations, participants who had been made accountable before the final decision tended to show a larger difference between the MAUT predictions for their chosen and for their average non-chosen alternative than participants who had been made accountable before the first decision ($M = 16.65$ vs. $M = 10.32$). This suggests that they made unidimensional value and weight judgements which supported their chosen over their average non-chosen alternative more strongly than participants who had already been made accountable before the first decision. The mean difference in MAUT predictions for the chosen and average non-chosen alternative after the final decision is shown in Figure 6.6.

The analysis on the *difference* between the first and the second elicitation revealed no effects which reached standard significance levels, but, again, there was a weak tendency for a significant main effect of time of accountability ($F(1,54) = 2.48$, $p = .121$), with participants who had been made accountable before the final decision showing a larger and, more importantly, positive difference compared to participants who had been made accountable before the first decision ($M = 4.90$ vs. $M = -1.00$). The mean value obtained for participants who had been made accountable before the final decision was significantly different from 0 ($t(27) = 2.30$, $p = .029$, two-sided).

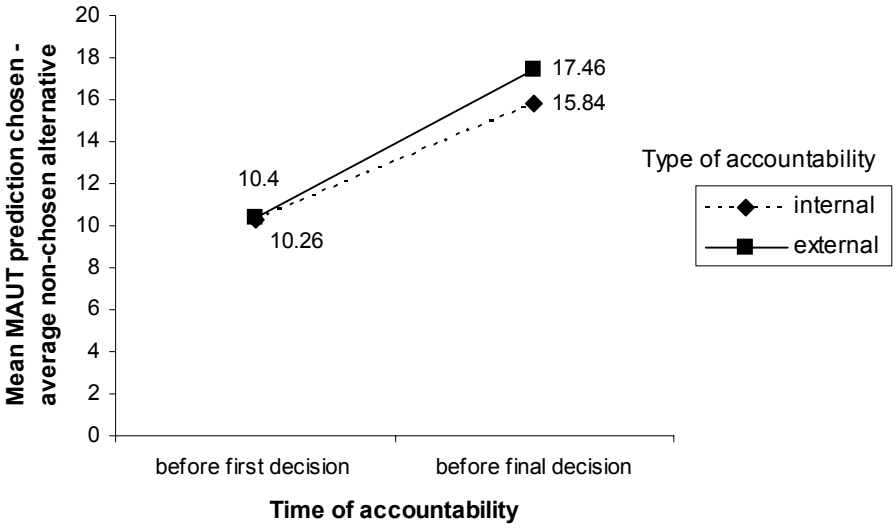


Figure 6.6. Mean MAUT predicted evaluative difference between the chosen and average non-chosen alternative after the final decision.

Since participants were encouraged to particularly differentiate between their chosen alternative and another, more or less randomly selected alternative (by telling them about a stalemate between these two alternatives, which their own decision would break), differentiation and consolidation processes were also analysed in terms of the difference between the chosen alternative and the competing alternative. In addition to the factors of time of accountability and type of accountability, the analyses included the preference rank of the competing alternative as a factor, that is, whether the competing alternative had been the second-best, third-best or fourth-best alternative at the first decision, based on the suitability ratings participants had provided then. Again, only participants who had not changed their choice after the first decision were included in the analysis.

3c. Ratings of the chosen alternative vs. the competing alternative

The overall mean difference between the rating for the chosen alternative and the rating for the competing alternative was 30.60 after the first decision and 20.60 after the final decision. This showed that, instead of increasing the judgmental difference between the chosen and the competing alternative, participants decreased it. The overall mean change

from the first to the final decision (M = -10.00) was significantly different from 0 ($t(57)$ = 4.59, p = .000, two-sided).

Three-factorial ANOVAs testing the effects of time of accountability, type of accountability and preference rank on the judgmental difference between the chosen and the competing alternative *after the first decision* revealed the expected highly significant main effect of preference rank ($F(2,46)$ = 15.86, p = .000). The higher the preference rank of the competing alternative, the smaller the difference between the chosen and the competing alternative (2nd rank: M = 15.54, 3rd rank: M = 32.49, 4th rank: M = 48.18). Post-hoc tests showed that this linear trend was highly significant (p = .000). No other effects were significant.

The ANOVA on the judgmental difference between the chosen alternative and the competing alternative *after the final decision* again showed a highly significant main effect of preference rank ($F(2,46)$ = 8.09, p = .001). As before, the lowest difference was obtained for the competing alternative that had been rated second-best after the first decision and highest for the competing alternative that had been rated fourth-best, with the third-best competing alternative in-between (2nd rank: M = 11.08, 3rd rank: M = 24.71, 4th rank: M = 29.94). The linear trend was again highly significant (p = .000). In addition, although the triple interaction between time of accountability, type of accountability and preference rank only showed a very weak tendency to be significant ($F(2,46)$ = 1.69, p = .196), simple main effect analyses showed that there was a significant interaction between time of accountability and type of accountability for participants whose competing alternative had been the least preferred at the first decision ($F(1,46)$ = 4.03, p = .050). In this case, when participants had been made accountable before the final decision, the obtained difference between the chosen and competing alternative was larger for externally accountable participants than for internally accountable participants (M = 38.75 vs. M = 23.33), whereas if participants had been made accountable before the first decision, the opposite was true (M = 37.50 vs. M = 22.33).

The main effect of preference rank was the only significant effect in the analysis on the *change* in judgmental difference between the chosen and the competing alternative from the first to the final decision ($F(2,46)$ = 3.94, p = .026). Again, the smallest difference between the second and first elicitation was obtained for participants whose competing

alternative had been their second-best at the first decision, and the highest difference was obtained for participants whose competing alternative had been their least preferred, with the third-best alternative in-between (2nd rank: *M* = -4.46, 3rd rank: *M* = -9.58, 4th rank: *M* =-18.24). The linear trend was highly significant (p = .007).

3d. MAUT predictions for the chosen alternative vs. the competing alternative.

The overall mean difference in MAUT predictions for the chosen alternative and the competing alternative at the first elicitation was 11.30, the overall mean difference at the second elicitation was 13.11. This meant that participants showed some differentiation and consolidation by slightly increasing this difference from the first to the second elicitation (overall mean: 1.81). This increase was not significant, however.

The ANOVAs performed on the difference in MAUT predictions between the chosen alternative and the competing alternative both at the *first elicitation* and at the *second elicitation* only revealed significant main effects of preference rank (1st elicitation: $F(2,46) = 10.05$, $p = .000$, 2nd elicitation: $F(2,46) = 8.23$, $p = .001$). As before, the difference between the chosen alternative and the competing alternative was smallest when the competing alternative had been the second-best at the first decision and largest when it had been the least preferred, with the third-best alternative in-between (1st elicitation: 2nd rank: *M* = 2.66, 3rd rank: *M* = 8.60, 4th rank: *M* = 26.21, 2nd elicitation: 2nd rank: *M* = 4.77, 3rd rank: *M* = 12.27, 4th rank: *M* = 25.73). In both cases, the linear trend was highly significant (p = .000).

The analysis of the *change* in these difference values from the first to the second elicitation did not reveal any significant effects. It had been expected that participants who had been made accountable before the final decision would show a larger, positive difference than participants who had been made accountable before the first decision, indicating stronger consolidation. Such a directional trend was apparent in the data. A positive difference was obtained for participants who had been made accountable before the final decision (*M* = 4.10), whereas for participants who had been made accountable before the first decision the obtained difference was negligible and slightly negative (*M* = -0.32). Because of considerable unsystematic variance in the data, this difference between conditions was not significant, however. Similarly, externally accountable participants tended to show a larger difference

between the second and first elicitation than internally accountable participants (M = 4.39 vs. M = -0.39), but, again, the difference between conditions was not significant.

Summary of the results for information evaluation and integration at the final decision

The results for the degree of bias in information evaluation and integration during the final decision proved largely inconclusive. It had been expected that participants who had been made accountable before the final decision and externally accountable participants would show more evidence of differentiation and consolidation processes compared to participants who had been made accountable before the first decision and internally accountable participants, respectively. Hardly any of the measures of differentiation and consolidation showed significant effects of either time or type of accountability, however. Participants' changes of attribute weights from the first elicitation, before any manipulation of accountability had occurred, to the second elicitation after the final decision failed to show the expected increase of weights for attributes on which the chosen alternative performed well and/or weights on which the non-chosen alternatives performed badly, but pointed to a regression effect instead; high weights were reduced, whereas low weights were increased. After controlling for any initial differences in weight judgements between conditions, no significant accountability effects were obtained. There was a marginally significant tendency for participants who had been made externally accountable before the first decision to increase the weight for their individually most and third-most important attribute, whereas participants who had been made externally accountable before the final decision decreased it. This was opposite to what had been expected. Possible explanations for the failure to obtain the expected effects will be discussed at the end of this chapter.

The analyses performed on actual and MAUT predicted overall ratings of alternatives also did not consistently support the hypotheses. Apart from a marginally significant effect of time of accountability on the difference between the MAUT predicted rating of the chosen and the average non-chosen alternative after the final decision, no significant effects of the manipulated variables were observed. The marginally significant effect of time of accountability suggested that participants who had been made accountable before the final decision, as expected, tended to show a larger evaluative difference between their chosen and

average non-chosen alternative than participants who had been made accountable before the first decision. Other data patterns also showed the expected trend. For example, participants who had been made externally accountable before the final decision showed signs of differentiation and consolidation by an increased difference in overall ratings between their chosen and average non-chosen alternative, whereas in all other conditions no change or even a decrease was observed.

Interestingly, whereas the actual ratings for the chosen compared to the competing alternative did not show any differentiation and consolidation effects (on the contrary, the evaluative difference between them was reduced in all conditions), the MAUT predictions of these ratings did; as expected, the largest increase in the difference between the chosen and competing alternative was found in the condition with external accountability introduced before the final decision and, more generally, when accountability had been introduced before the final decision compared to before the first decision. A possible reason for this dissociation may be that the fact that the competing alternative had been said to be part of a stalemate, prevented participants from openly devaluing it. However, when devaluation was less obvious and more subtle, participants took the chance to engage in it. Consistent with this explanation is the fact that 26% of the participants whose data was included in the analysis adopted the competing alternative as their second-best alternative when making their final decision, thereby expressing some conformity with and acknowledgement of the preferences of the other committee members. This behaviour also meant that there was little scope to increase the difference between the chosen and the competing alternative further, unless participants had chosen to rate all alternatives other than the one they preferred rather badly. Furthermore, the observation that participants who had been made externally accountable before the final decision showed the largest tendency to minimise the difference between the overall rating for their chosen and the competing alternative supports the PMA prediction that their motivation to attend to external preferences should have been most salient.

Justifications

Participants' written justifications were content-analysed. The unit of analysis was an argument, defined as the smallest unit that provided some meaning in the experimental context. Two independent judges sorted the arguments with respect to the categories listed

below. Inter-judge correlations were moderate to high (range: r = .64 to r = .91, mean: r = .87). Differences in categorisation were resolved by discussion. The results are presented in Table 6.11.

Table 6.11. Justification results (percentages in brackets).

	Time of Accountability						
	Before First Decision		Before Final Decision		*Significance of Effects*		
Type of Accountability	Internal	External	Internal	External	*Type*	*Time*	*Type x Time*
1. Type of argument							
direct	2.17 (46)	3.22 (59)	3.50 (70)	4.05 (69)		*	
elaboration	2.83 (49)	2.00 (33)	1.33 (27)	1.78 (28)		(*)	
false reproduction	0.33 (5)	0.39 (8)	0.17 (3)	0.17 (2)		(*)	
total	5.33 (100)	5.61 (100)	5.00 (100)	6.00 (100)			
2a. Consonant arguments							
positive aspects of the chosen alternative	4.56 (79)	5.06 (90)	4.22 (86)	4.72 (84)			
negative aspects of the non-chosen alt.	0.28 (9)	0.05 (1)	0.11 (2)	0.56 (7)			*
total	4.84 (88)	5.11 (91)	4.33 (88)	5.28 (91)			
2b. Dissonant arguments							
positive aspects of the non-chosen alt.	0.00 (0)	0.28 (5)	0.11 (2)	0.44 (5)	(*)		
negative aspects of the chosen alt.	0.50 (12)	0.22 (4)	0.56 (10)	0.28 (4)	*		
total	0.50 (12)	0.50 (90)	0.67 (12)	0.72 (9)			
3. Complexity	0.15	0.14	0.10	0.11			

Note: Discrepancies are due to rounding errors.
(*) $p \leq .10$
* $p \leq .05$

The following types of dependent variables were analysed:

- Total number of arguments

  This number included three subtypes of arguments: (1) direct arguments, which were directly related to information presented during the experiment, (2) elaborations, that is,

arguments in which participants had elaborated on information presented during the experiment and which could not be directly inferred from this information[2] , and (3) false reproductions of information, where participants' arguments contained incorrect information.

- Consonant vs. dissonant arguments

  Consonant arguments were arguments which mentioned a positive aspect of the chosen alternative or a negative aspect of a non-chosen alternative, whereas dissonant arguments were arguments which mentioned a positive aspect of a non-chosen alternative or a negative aspect of the chosen alternative.

- Complexity of argumentation

  The complexity of argumentation was measured by the ratio of the number of two-sided arguments to the sum of one-sided and two-sided arguments. Two-sided arguments were defined as arguments which mentioned two alternatives, comparing them in some way, whereas one-sided arguments were arguments that only mentioned one alternative. The complexity index ranged between 0 and 1, where 0 indicated that only one-sided arguments had been presented and the complexity of argumentation was low, and 1 indicated that only two-sided arguments had been presented and the complexity of argumentation was high. It was expected that participants who had only been made accountable after they had already made a decision, that is, participants who had been made accountable before the final decision, would show a higher complexity of argumentation than participants who had already been made accountable before the first decision.

1. Type of arguments

The mean total number of arguments listed by participants was 5.49. Of these, 3.24 were arguments directly related to information presented during the experiment, 1.99 were elaborations and 0.26 were false reproductions.

[2] For example, a participant may have inferred from the information that the job candidate's thesis topic had not been related to the research interests of his or her future PhD supervisor that this would broaden the focus of the team.

Two-factorial ANOVAs testing the effects of type of accountability and time of accountability were performed on these variables. For the total number of arguments, no significant effects of the experimental variables were obtained. Nevertheless, in order to eliminate the effect of the total number of arguments mentioned, the different types of arguments were transformed into percentages and these percentages were analysed. For direct arguments, a significant main effect of time of accountability was observed ($F(1,68) = 6.36$, $p = .014$). Participants who had been made accountable before the final decision mentioned a higher percentage of direct arguments than participants who had been made accountable before the first decision ($M = 70\%$ vs. $M = 53\%$). In the ANOVA performed on the percentage of elaborations, the main effect of time of accountability was marginally significant ($F(1,68) = 3.88$, $p = .053$). Participants who had been made accountable before the final decision tended to present a smaller percentage of elaborated arguments than participants who had been made accountable before the first decision ($M = 28\%$ vs. $M = 41\%$). Finally, a marginally significant main effect of time of accountability was also obtained for the number of false reproductions ($F(1,68) = 3.39$, $p = .070$). The percentage of false reproductions tended to be larger for participants who had been made accountable before the first decision than participants who had been made accountable before the final decision ($M = 7\%$ vs. $M = 3\%$).

2a. Consonant arguments

The overall mean total number of consonant arguments was 4.89, of which 4.64 mentioned positive aspects of the chosen alternative and 0.25 mentioned negative aspects of the non-chosen alternatives. The difference between the latter two values was highly significant ($t(71) = 13.80$, $p = .000$), which suggests that when they presented consonant arguments, participants showed a significantly stronger overall tendency to focus on the advantages of the chosen alternative than the disadvantages of the non-chosen alternative.

Two-factorial ANOVAs were conducted to test the effects of type of accountability and time of accountability on the percentages of the two different subtypes of consonant arguments as well as the percentage of consonant arguments overall, relative to the total number of arguments. The analysis for the percentage of consonant arguments overall did not reveal any significant effects nor did the analysis for the percentage of positive aspects

mentioned for the chosen alternative. However, a significant interaction between time of accountability and type of accountability was observed in the analysis of the percentage of negative aspects mentioned for the non-chosen alternatives ($F(1,68) = 4.91$, $p = .030$). Simple main effects analyses indicated that there was a significant 'type of accountability' effect for participants who had been made accountable before the first decision only ($F(1,68) = 4.03$, $p = .049$). Internally accountable participants who had been made accountable before the first decision produced a higher percentage of arguments that stressed the negative aspects of the non-chosen alternatives than participants who had been made externally accountable before the first decision ($M = 9\%$ vs. $M = 1\%$). An opposite tendency was observed for participants who had been made accountable before the final decision ($M = 2\%$ vs. $M = 7\%$), this was not significant, however.

2b. Dissonant arguments

The overall mean total number of dissonant arguments was 0.60, of which 0.21 mentioned positive aspects of the non-chosen alternatives and 0.39 mentioned negative aspects of the chosen alternative. The difference between the latter two values was not significant. Dissonant arguments were given significantly less frequently than consonant arguments, however ($M = 0.60$ vs. $M = 4.64$, $t(71) = 13.47$, $p = .000$).

Again, two-factorial ANOVAs were employed to test the effects of type of accountability and time of accountability on the dependent measures. For the percentage of dissonant arguments overall, no significant effects were obtained. However, both the percentage of arguments mentioning positive aspects of the non-chosen alternatives and the percentage of arguments mentioning negative aspects of the chosen alternative were affected by type of accountability. Externally accountable participants tended to mention a higher percentage of arguments focusing on the positive aspects of the non-chosen alternatives ($F(1,68) = 3.74$, $p = .057$, $M = 5\%$ vs. $M = 1\%$) and mentioned a significantly lower percentage of arguments focusing on the negative aspects of the chosen alternative ($F(1,68) = 5.91$, $p = .018$, $M = 4\%$ vs. $M = 11\%$) than internally accountable participants. These effects are shown in Figure 6.7 and Figure 6.8.

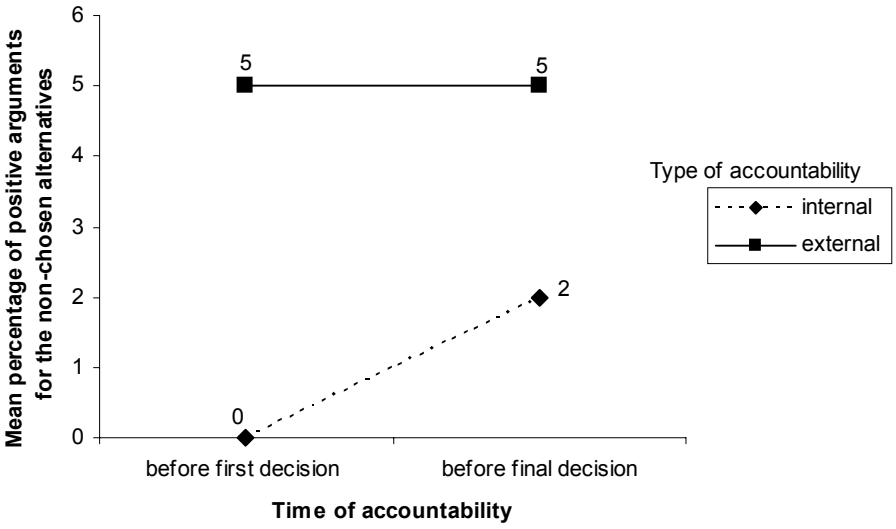


Figure 6.7. Mean percentage of arguments mentioning positive aspects of the non-chosen alternatives, as a function of type of accountability and time of accountability.

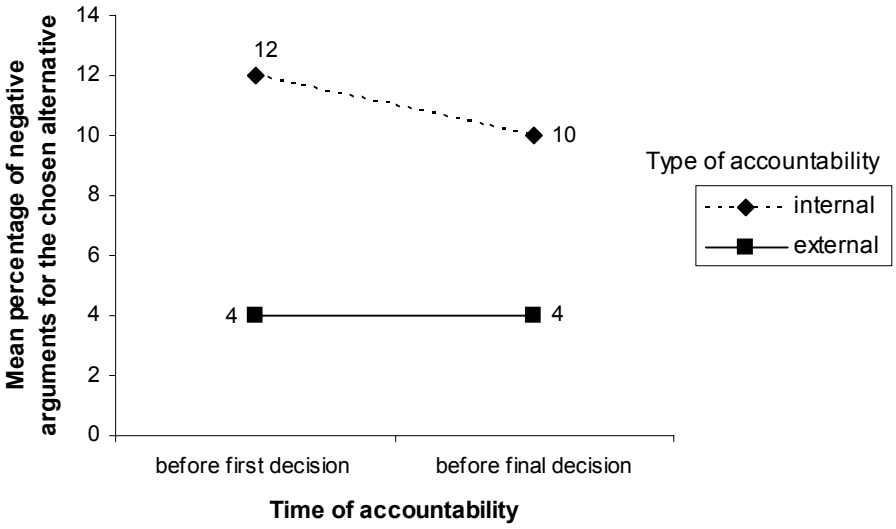


Figure 6.8. Mean percentage of arguments mentioning negative aspects of the chosen alternative, as a function of type of accountability and time of accountability.

3. Complexity of argumentation

The overall mean complexity index was 0.12 (maximum value 1), suggesting a generally very low complexity of argumentation.

Unlike expected, the complexity index was lower for participants who had been made accountable before the final decision than participants who had been made accountable before the first decision (M = 0.10 vs. M = 0.14). This difference was not significant, however. The main effect of type of accountability and the interaction between time of accountability and type of accountability were also not significant.

Summary of the justification results

Although expected, the analyses of participants' justifications did not reveal any significant effects of time of accountability on depth and complexity of argumentation. Participants who had been made accountable immediately before the final decision did not produce more arguments overall than those who had been made accountable before the first decision, nor were the justifications of participants who had been made accountable before the final decision more complex than those of participants who had been made accountable before the first decision. With regard to complexity, if anything, a trend in the opposite direction could be observed.

Time of accountability did, however, affect the relative frequency of particular types of arguments. Notable is the higher percentage of elaborations and lower percentage of direct arguments produced by participants who had been made accountable before the first rather than the final decision. This is likely to be due to the fact that these participants had had a longer time span between first hearing that they were going to be accountable and giving their justifications than participants who only got to know before the final decision that they would have to justify their decision. Somewhat surprising therefore is the higher percentage of false reproductions for participants who had been made accountable before the first rather than the final decision, because they should have been less likely to forget arguments, given that they knew from the start that they would have to justify their decision. Hence, it is likely that false reproductions were used in a strategic manner to support the chosen alternative. Indeed, a closer inspection of the particular type of false reproductions created shows that a significant

effect of time of accountability is only obtained for false reproductions that favour the chosen alternative ($F(1,68) = 4.31$, $p = .042$, accountability before the first decision: $M = 6\%$ vs. $M = 2\%$). Also, it is interesting to note that false reproductions favouring the non-chosen alternatives or devaluing the chosen alternative did not occur at all.

Type of accountability seemed to particularly affect the number of dissonant arguments presented. Externally accountable participants produced relatively more arguments mentioning positive aspects of the non-chosen alternatives than internally accountable participants. Interestingly, when it came to mentioning negative aspects of the chosen alternative, internally accountable participants did so significantly more often than externally accountable participants. This pattern of results is consistent with external accountability inducing concerns about external evaluation. Externally accountable participants, particularly those who had been made accountable before the final decision, presumably were acutely aware of the fact that some of the people they had to justify their choice to would be opposed to it and, as part of their strategy, seem to have mentioned positive aspects of their non-chosen alternatives, showing that they had given some thought to them, but then to have refuted these positive aspects. The latter assumption receives some support by the significant interaction effect between time and type of accountability observed for the percentage of negative aspects of the non-chosen alternatives. Although the interaction effect was caused by the large difference between internal and external accountability before the first decision, as expected, participants who had been made externally accountable before the final decision mentioned a higher percentage of negative aspects of the non-chosen alternatives than participants who had been internally accountable. However, given that the number of arguments mentioning the non-chosen alternatives was generally very small, these results must be viewed with some caution.

Ancillary measures

1. Difficulty of the first decision

After their first decision, participants had been asked to rate the difficulty of this decision on a scale from 1 (very easy) to 5 (very difficult). No significant differences emerged between conditions (no accountability: $M = 3.14$, internal accountability: $M = 3.06$, external

accountability: $M = 3.28$). As expected, however, the highest rating was observed for participants who had been made externally accountable.

2. Certainty about having made a good first decision

Participants' responses to the question about how certain they were to have made a good first decision, given on a scale from 1 (very uncertain) to 5 (very certain), showed a significant difference between conditions in the expected direction. Participants who had been made accountable before making this decision (no matter whether internally or externally) felt significantly less sure to have made a good decision than participants who had not been made accountable ($M = 2.50$ vs. $M = 2.94$, $t(70) = 2.35$, $p = .021$). There was no significant difference between the 'internal accountability' and the 'external accountability' conditions ($M = 2.44$ vs. $M = 2.56$).

3. Difficulty of the justification

After the written justification of their final decision, participants had rated how difficult this justification had been. A 2 (type of accountability: internal vs. external) by 2 (time of accountability: before the first decision vs. before the final decision) ANOVA yielded a marginally significant main effect of type of accountability ($F(1,68) = 3.00$, $p = .088$). As expected, participants in the 'external accountability' conditions tended to find it more difficult to justify their decision than participants in the 'internal accountability' conditions ($M = 2.89$ vs. $M = 2.56$). No other significant differences emerged.

4. Expectation of having to justify one's decision

At the end of the experiment, participants had rated the extent to which they had expected to have to justify their decision on a scale from 1 (not at all) to 5 (very much). A two-factorial ANOVA testing the effects of type of accountability and time of accountability on participants' responses yielded a significant main effect of time of accountability ($F(1,68) = 4.73$, $p = .033$). Participants who had been made accountable before the final decision expected to have to justify their decision more strongly than participants who had been made

accountable before the first decision (M = 3.11 vs. M = 2.53). No other effects were significant.

5. Need to make a good decision

For the reported need to make a good decision, rated on a scale from 1 (very weak) to 5 (very strong), the ANOVA analysis revealed a marginally significant main effect of time of accountability ($F(1,68) = 2.97$, $p = .089$). As expected, participants who had been made accountable before the final decision tended to report a greater need to make a good decision than participants who had been made accountable before the first decision (M = 2.08 vs. M = 1.83). There were no other significant effects.

6. Need to be able to justify the first and final decision

The analyses of the responses to the questions about how strong participants' perceived need was to be able to justify their decision to themselves and to others at the first and at the final decision did not reveal the expected pattern. It had been expected that participants in the 'external accountability before the first decision' condition would report a greater need of being able to justify their decision to others at the first decision than participants in the 'internal accountability before the first decision' condition. No difference had been expected between the responses of these participants to the question about their need to be able to justify their decision to themselves at that time. Also, it had been expected that participants with accountability introduced before the first decision would generally report a greater need to justify their first decision and participants with accountability introduced before the final decision a greater need to justify their final decision.

The two-factorial ANOVAs performed on these measures revealed no significant effects, except for a main effect of time of accountability on participants' need to justify their final decision to themselves ($F(1,68) = 4.86$, p =.031). Participants who had been made accountable before the first decision reported a greater need to justify their final decision to themselves than participants who had been made accountable before the final decision (M = 4.22 vs. M = 3.81).

A repeated measures ANOVA was also conducted which, apart from the two between-subjects factors of type of accountability and time of accountability, included two within-participants factors, type of rating (need to justify to self vs. to others) and decision (need to justify the first decision vs. the final decision). The analysis revealed a highly significant main effect of type of rating ($F(1,68) = 8.41$, $p = .005$). Participants reported a stronger need to be able to justify their decisions to themselves than to others ($M = 4.01$ vs. $M = 3.65$). There was also a marginally significant interaction between type of rating and decision ($F(1,68) = 3.08$, $p = .084$). Whereas there was no difference between the reported need to be able to justify the first and the final decision where justification to the self was concerned ($M = 4.01$ vs. $M = 4.01$), participants tended to report a stronger need to be able to justify their final decision to others than their first decision ($M = 3.75$ vs. $M = 3.54$). This effect was further qualified by a marginally significant triple interaction between type of rating, decision and type of accountability ($F(1,68) = 3.95$, $p = .051$). The previously described interaction pattern between type of rating and decision could be observed for participants who had been made internally accountable, participants who had been made externally accountable, however, did not show this effect.

A correlational analysis indicated that the correlation between the perceived need to justify the first and the final decision to others was higher for participants who had been made accountable before the first decision than for participants who had been made accountable before the final decision (r (36)= .72, $p = .000$ vs. r (36) = .58, p = .000). This is consistent with expectations, assuming that participants who had been made accountable before the first decision had a strong need to justify the first decision, which remained high for the final decision, whereas participants who had been made accountable before the final decision would only perceive a strong need to justify the final but not the first decision. The same result was not obtained, however, for the correlation between the perceived need to justify the first and the final decision to themselves; contrary to what would have been predicted, this was higher for participants who had been made accountable before the final compared to the first decision (r (36)= .53, $p = .000$ vs. r (36) = .35, p = .037).

Summary of the results for the ancillary measures

The analyses of the ancillary measures supported some assumptions about the effects of the experimental manipulations, but failed to support others. As expected, participants who had been made accountable prior to their first decision were less sure that they had made a good first decision than participants who had not. There were no significant differences in terms of the perceived difficulty of this decision, but, consistent with expectations, participants who had been made externally accountable expressed the greatest difficulty. As expected, they also found it more difficult to justify their decision than participants who had been made internally accountable. When asked about their need to make a good decision at the end of the experiment, participants who had only been made accountable before the final decision tended to report a greater need than participants who had already been made accountable before the first decision, suggesting, not surprisingly, that the most recent manipulation of accountability was still strongest in participants' minds. This was also borne out by responses to the question of how strongly participants had expected to justify their decision. Although there should not have been any differences once all participants had been made accountable in some way, participants who had been made accountable before the final decision reported a stronger expectation of having to justify their decision than participants who had been made accountable before the first decision.

The results that departed most strongly from the original predictions were the results relating to the need to be able to justify the first and the final decision to the self and others. It was not the case that participants who had been made externally accountable before the first decision reported a greater need to justify their first decision to others than participants who had been made internally accountable at that time, nor did participants who had been made accountable before the first decision generally report a greater need to justify their first decision, and participants who had been made accountable before the final decision a greater need to justify the final decision. Instead, participants generally expressed a stronger need to justify their decision to themselves than to others, and this tendency was stronger for participants who had been made accountable before the first compared to the final decision. These effects may have been a consequence of the external accountability manipulation not being strong enough, particularly not after some time had passed, and/or impression management concerns on the part of the participants, who may have wanted to signal to the

experimenter that they were taking the task seriously. The fact that participants tended to report a stronger need to justify their final rather than their first decision to others supports this claim. Also, asking participants to make these kinds of judgements retrospectively, may have biased them towards the most recent and therefore most salient instruction.

Discussion and conclusions

The aim of the present study was to validate the PMA by investigating the effects of internal and external accountability demands and the time when these demands were introduced (before making a first vs. after making a first and before making a final decision). The central assumption of the PMA is that accountability both enhances the accuracy of and the degree of bias in information processing when arriving at a decision, but that the relative strength of these effects depends on the extent to which the manipulation makes salient the decision maker's personal responsibility for the consequences of his or her decision and the expectation of being evaluated by an external audience. The PMA predicts that becoming aware of being personally responsible for one's actions will enhance the motivation to be accurate, which, in turn, will increase the depth of information search and complexity of information integration. Evaluation apprehension, on the other hand, will enhance the motivation to make a decision that can be easily justified to one's evaluators, which, in turn, will increase a bias in information evaluation and integration towards the perceived norms of the external audience - if those are salient - or towards letting the preferred alternative appear maximally superior to the others. Since internal accountability only makes personal responsibility salient, without drawing the decision maker's attention to any external norms, whereas the introduction of external accountability demands usually involves both personal responsibility and evaluation cues, a comparison of the effects of the two should show a deeper and more complex information processing both in conditions of internal and external accountability, compared to a control group without accountability demands, but a more biased information processing for externally accountable participants compared to participants in the control group and 'internal accountability' condition. However, this difference between the effects of internal and external accountability with respect to the degree of bias in information evaluation and integration, according to the PMA, should be diminished if individuals are only made accountable once they have already committed

themselves to an alternative, because then both internally and externally accountable individuals are assumed to have a motivation to bias their information processing towards supporting the alternative they have previously chosen. In the case of externally accountable participants, this tendency may be particularly strong if they perceive a conflict between their own choice and the choice preferred by their external audience.

The present study therefore manipulated internal and external accountability either before a first decision or before a final decision and monitored participants' information search, evaluation and integration of information during the decision making process. Internal accountability was induced by making personal responsibility for the decision salient while keeping the decision anonymous at the same time, thereby minimising external evaluation apprehension. This was achieved through a committee scenario. The external accountability manipulation made both personal responsibility and external evaluation salient. The findings lent some support to the PMA predictions, but also highlighted methodological problems. In particular, features of the experimental task seemed to have occasionally prevented expected effects from occurring.

Information search before the first decision

Unlike expected, participants who had been made accountable before the first decision did not differ significantly from non-accountable participants in terms of depth and variability of their information search, although the observed trends were usually in the right direction; the information search of participants who had been made accountable tended to be deeper, less variable and more compensatory than that of participants who had not been made accountable. Apart from the fact that there was a large individual variability in information search which may have prevented some effects from becoming statistically significant, the generally very exhaustive information search may have masked any differences between conditions. This exhaustive search, even for participants who were not accountable, may have been caused by telling participants before they made their decision that they additionally would have to judge the quality of all alternatives. As Billings and Scherer (1988) and Koele and Westenberg (1992) have previously pointed out, the response mode of judgement typically results in a deeper and more compensatory information processing than that of choice. Surprisingly, pre-decisional external accountability (and not the absence of

accountability) resulted in the lowest total number of items searched. It may be speculated that the prospect of having to justify one's decision to others enhances memory for the searched information and therefore results in fewer repeated searches. Consistent with this assumption, the percentage of repeatedly searched information was indeed lowest in the 'external accountability' group, although not significantly so. Finally, participants under pre-decisional external accountability did not show a significantly higher percentage of information searched for the individually most important attribute than participants in the other conditions.

Information evaluation and integration at the first decision

The findings concerning information evaluation and integration during the first decision, like the findings for the first information search, were largely inconclusive, due to the absence of any significant effects. However, not only were there no significant effects, but there were also some trends in the data that were inconsistent with the pattern predicted by the PMA. It had been expected that externally accountable participants would be particularly likely to choose the alternative with the best outcome on the individually most important attribute, and that they would show the largest evaluative differentiation between their chosen and the other alternatives, compared to conditions of no and internal accountability. Generally, the frequency with which the alternative with the best outcome on the individually most important attribute was chosen was quite high (60%), but externally accountable participants were no more likely to do so than participants in the other conditions. The general popularity of such a lexicographic decision rule may be explained by the relatively high complexity of the decision situation (4 alternatives that differed on 5 attributes), which typically results in the employment of non-compensatory decision strategies (see Payne, Bettman, & Johnson, 1993, for a review). Also, instead of resulting in the largest, external accountability resulted in the smallest observed difference between the rating of the overall quality of the chosen alternative and the average rating for the non-chosen alternatives. This difference was largest for participants under internal accountability. The main effect of type of accountability was not significant, but the same trend was observed for the actually observed difference between these ratings and the difference predicted by a linear MAUT model, suggesting that external accountability resulted in a decreased rather than increased

differentiation. This may reflect participants' uncertainty in terms of which alternative would be preferred by the other committee members to which they had been made accountable. As a result of this normative uncertainty, they may have resorted to a strategy of 'sitting on the fence', whereby they kept the difference between their preferred and the other alternatives small in order to be able to endorse a different alternative that appeared more justifiable, should it be necessary.

Information search before the final decision

Analysis of information processing exhibited during the final decision again only provided partial support for the PMA. It had been expected that participants who had been made accountable immediately before the final information search would exhibit a deeper, less variable and more compensatory search than participants who had already been made accountable before the first decision. This was not generally observed (except in the lower variability of search across attributes for different cells searched, which was shown by participants who had been made accountable before the final compared to the first decision), but, again, features of the task participants were asked to perform may help to explain why. Before they engaged in the process of arriving at a final decision, all participants were told that a stalemate had arisen between the other committee members regarding which alternative they supported. One of these stalemate alternatives was said to be the one participants had expressed an initial preference for, the other was allocated in such a way that the number of particular combinations of alternatives presented did not differ largely between conditions. The results for the final information search suggested that participants did not consider all alternatives anymore, but confined their search to a subset of the available information instead, namely particularly information about those alternatives between which the stalemate had been said to have arisen. This effect tended to be stronger when accountability had been introduced before the final decision than when it had been introduced before the first decision, and particularly strong in the ‘internal accountability before the final decision' condition, especially in terms of the concentration of search on the chosen alternative. Since participants in this condition did not expect to have to justify their decision to the other committee members, they were free to concentrate on the two alternatives that were presented to them and check whether the alternative which they had previously chosen was still better than the

one allegedly preferred by two of the other committee members. Participants who had been made externally accountable before the final decision, on the other hand, expected that they would have to justify their decision to the other committee members and therefore, in addition to searching the two presented alternatives intensely, seem to have looked for comparison information about other alternatives, in order to be well-prepared for a discussion. The finding that internal accountability introduced before the first decision led to a particularly thorough and compensatory search of all alternatives cannot be easily explained. It may be that, since participants in this condition were only accountable to themselves and their accountability was not immediately salient, there was less motivation for them to attend to the alleged opinions of other committee members and it was more important to make sure that of all available alternatives the one they had chosen before was really the best one.

With hindsight, it appears that the procedure of presenting participants with a stalemate between two alternatives (which had been adopted in order to have a good cover story for a final decision) was problematical, because it focused participants' attention on the two stalemate alternatives and thereby made external norms salient, not only for participants who had been made externally accountable but also for those who had been made internally accountable. This may explain why the responses to the questions about how strongly participants had felt a need to justify their first and final decision to themselves and others did not show the expected pattern, but expressed a generally stronger need to justify the final decision to others, regardless of condition.

The analyses on the type of information searched before making a final decision revealed a tendency for externally accountable participants to search a lower percentage of information for the attribute they individually found least important than internally accountable participants. This indirectly supports the idea that individuals who have been made externally accountable pay more attention to information they consider important in a given decision situation, to the detriment of less important attributes. When looking at attributes as such rather than their importance rank, it emerged that externally accountable participants also tended to search a lower percentage of information for the 'student representative on university committees' attribute than participants in the 'internal accountability' conditions. Given that this attribute had received the lowest mean importance rating of all attributes in information set 2, this finding is consistent with the earlier

observation for the individually least important attribute. However, another effect may have been at work here as well. Participants were informed that the other committee members were junior lecturers and professors. Anticipating having to justify their decision to them may have made participants reluctant to consider information that was clearly related to student issues and as such may not have been regarded as convincing by the other committee members. This explanation would be consistent with Hsee's (1995, 1996) model of elastic justification, which suggests that tempting attributes are only taken into account when individuals perceive the freedom to do so without negative repercussions.

An important feature of the second information set was that its information was largely redundant with the information contained in information set 1, in order to give participants the opportunity to anticipate the outcomes of their search when they accessed the cells of the information board. This made it possible to test the extent of confirmation bias in participants' information search, that is, a bias to support the chosen and/or devalue the non-chosen alternatives. Although the measure developed to capture confirmation bias, the Biased Search Index (BSI), surprisingly did neither indicate a significant overall bias to support the chosen nor a bias to devalue the other stalemate (competing) and average non-chosen alternative, the overall mean BSI ratio for the chosen to the average non-chosen alternative did suggest that participants' information search was biased to support the chosen over the average non-chosen alternative. This was not true, however, for the chosen compared to the competing alternative, probably because the final information search was strongly and equally focused on the two stalemate alternatives. Whereas the expected stronger bias in the post-decisional compared to the pre-decisional accountability conditions could not be observed, there were a number of significant effects of type of accountability and significant interactions between time and type of accountability on the amount of bias displayed. Externally accountable participants showed a significant tendency to search more information that they expected to support the chosen compared to the average non-chosen alternative, and they did significantly more so than internally accountable participants. They also showed a significantly stronger bias to support the chosen over the competing alternative than participants in the 'internal accountability' conditions. The cause of these effects may be found in the fact that externally accountable participants experienced the conflict between their own choice and that favoured by some other members of the committee more strongly, given they

anticipated a public justification of their choice, and therefore may have felt a need to bolster their chosen alternative more than internally accountable participants. At the same time, they also seem to have looked for the advantages of the other alternatives, in case they would be challenged about them. The latter is suggested by the fact that participants who had been made externally accountable before the final decision showed a significantly weaker bias to search information that devalued the average non-chosen alternative and even a slight bias towards searching information that supported the average non-chosen alternative compared to participants who had been made externally accountable before the first decision. The problem that was mentioned earlier, namely that informing participants about a stalemate between two alternatives before they made their final decision made external norms salient, not only for externally accountable but also for internally accountable participants, is underlined by the fact that participants who had been made internally accountable before the final decision searched more information that devalued the competing alternative compared to participants who had been made internally accountable already before the first decision, and showed the strongest such bias of all conditions.

Although the analyses that tested confirmation bias revealed some significant and interesting effects, it must be mentioned here that the concept of selective exposure or confirmation bias in information search and its measurement is fraught with problems, as already discussed in Chapter 4. Support for its existence has been scarce, despite a large number of studies within the tradition of dissonance theory. Canon's (1964) and Freedman's (1965b) studies, in particular, indicated that external accountability (operationalised by telling participants that they would have to engage in a debate or would be presented with opposing arguments) resulted in a preference for dissonant over consonant information, as the former was perceived to be more useful for the task. Similarly, in the present study, participants who had been made externally accountable before the final decision displayed the strongest tendency to search information that supported the average non-chosen alternative.

Also, a particular problem of selective exposure paradigms has been the interpretation of attention to particular information. While increased attention to dissonant information may signal a preference for it, it may also signal the opposite, namely an attempt to refute the unwanted information and as such can be seen as an attempt to reduce dissonance in the long term and bolster the chosen alternative (Grabitz & Grabitz-Gniech, 1973; Wyer & Frey,

1983). Ditto and Lopez (1992) have indeed demonstrated that information inconsistent with a preferred conclusion is examined more critically than information consistent with a preferred conclusion and as a consequence, more information is required to reach a preference-inconsistent conclusion than to reach a preference-consistent one. The present results, although encouraging, must therefore be interpreted with caution; the fact that in some conditions less consonant and more dissonant information was searched, does not necessarily mean that there was no attempt to support the chosen alternative.

Finally, although an attempt was made to create information for the second information set that could be inferred from the information contained in the first set, the redundancy was not perfect. Some of the parallel attributes used were only judged as moderately similar on average, for example, the attributes 'job as a student research assistant' and 'computing skills'. While it is relatively safe to assume that a person who has been a student research assistant has good computing skills, otherwise they would not have got the job, it does not necessarily mean that someone who has not been a student research assistant before will not have good computing skills. Therefore, it cannot be ruled out that some outcomes had been anticipated by participants differently than intended and came as a surprise to them. In a future experiment employing the same procedure, it would be desirable to try to achieve real redundancy by using the same information, just expressed in a slightly different way.

Information evaluation and integration at the final decision

Analysis of participants' information evaluation during and after the final decision mainly focused on whether any differentiation and consolidation effects could be observed and whether they differed as a function of time and type of accountability. Despite the fact that participants were encouraged to adhere to their initial choice, by implying that two other committee members also preferred the candidate they preferred, a substantial number of them (19%) changed their decision. The likelihood of change did not differ as a function of type or time of accountability, however. It cannot be ruled out that some participants misunderstood the instructions and thought that all other committee members preferred a different alternative. Because any differentiation and consolidation processes would only be evident if participants had not changed their choice from the first to the final decision, the data for

participants for whom this was not the case were excluded from the analysis. This reduced the number of participants to 14 to 16 per condition and resulted in a loss of power. Hence, it may not be seen as surprising that many of the observed effects failed to reach standard significance levels.

Importance differentiation was assessed by calculating the difference between participants' normalised weights for the attributes of information set1 elicited before any of the experimental manipulations and after the final decision. It had been expected that participants would increase the weights for attributes they had initially considered important and on which their chosen alternative had good outcomes, and /or decrease the weight for attributes they had initially considered less important and on which their chosen alternative had bad outcomes. It had also been expected that this tendency would be stronger for participants who had been made accountable before the final rather than the first decision. The results did not provide any evidence for this; on the contrary, they indicated a regression effect, where the weight of important attributes was decreased and that of unimportant ones increased, resulting in a more equal weight distribution. However, accountability has been previously found to increase the dilution effect (Tetlock & Boettger, 1989; Tetlock, Lerner, & Boettger, 1996), where individuals inappropriately take into account, that is, assign a too high weight to non-diagnostic information, and it is possible that the participants in this study here also displayed this tendency, especially in the situation of normative ambiguity they were faced with. Once any differences between conditions with regard to the first weight judgements were kept constant, there were no significant effects of the experimental manipulations on the calculated measures, except for a marginally significant tendency for participants who had been made externally accountable before the first decision to increase the weight for their individually third-most important attribute, whereas participants who had been made externally accountable before the final decision decreased it. This was not in line with expectations, but could be interpreted as supporting the dilution argument made earlier. Participants made externally accountable before the final decision may be seen as particularly prone to showing such a tendency in their responses, as their motivation to be seen to take all information into account should be particularly strong.

Differentiation and consolidation effects had also been expected for actual and MAUT predicted overall ratings of alternatives. The only effect that approached significance was a

tendency for participants who had been made accountable before the final decision to show the expected larger MAUT predicted evaluative difference between their chosen and average non-chosen alternative compared to participants who had been made accountable before the first decision. However, also the fact that only participants who had been made externally accountable before the final decision showed signs of differentiation and consolidation, by increasing the difference between their overall rating for the chosen and their average non-chosen alternative, whereas in all other conditions no change or a decrease was observed, was consistent with predictions.

In addition, a very interesting dissociation between the results for the actually observed and the MAUT predicted evaluative difference between the overall ratings of the chosen and competing alternative was found. MAUT predictions were based on participants' unidimensional value and weight judgements before the experimental manipulations and after the final decision, and were calculated to test any joint effects of attractiveness restructuring and importance differentiation. The rationale behind this was that the MAUT prediction of the overall value of an alternative combines unidimensional values and weights in a multiplicatory fashion and thereby embodies the principle that attractiveness restructuring is most effective for attributes that are given a high weight. If significant attractiveness restructuring and importance differentiation occurred, the MAUT prediction based on unidimensional values and weights elicited after the final decision should be significantly larger than the MAUT prediction based on the first elicitation of unidimensional values and weights. It was found that, whereas the actual ratings for the chosen compared to the competing alternative did not show any differentiation and consolidation effects (on the contrary, the evaluative difference between these two alternatives was reduced over time in all conditions), the MAUT predictions of these ratings did; as expected, the largest (albeit not significant) increase in the difference between the chosen and competing alternative was found in the condition with external accountability introduced before the final decision and, more generally, when accountability had been introduced before the final rather than the first decision. It may be speculated that only when the opportunity to bolster the chosen alternative was subtle and covert, as in the case of unidimensional value and weights judgements, would participants engage in this process. When it was obvious, on the other hand, as in the case of having to rate the overall quality of their chosen and competing alternative, participants may

have felt reluctant to openly devalue the competing alternative, especially when they had been made externally accountable before the final decision. Another factor that may have meant that it was difficult to obtain the predicted differentiation and consolidation effects was the fact that a considerable number of participants (26% of those whose data had been entered in to the analysis) made the competing alternative their new second-best alternative, thereby acknowledging the alleged preferences of the other committee members. There were no significant systematic differences between conditions as to the likelihood of this happening, but it nevertheless meant that there was little scope for an increase in the difference between the chosen and the competing alternative; rather, the opposite was likely, given that the competing alternative had been judged worse before. These problems could have been avoided if the stalemate had always been said to involve participants' preferred and second-best alternative, as this would have eliminated the possibility for participants to keep their initial choice but change their second-best alternative. This would have forced them to decide between the two and would have made differentiation and consolidation processes more pronounced. It would not, however, have solved the problem of making different external norms salient to different participants.

Although studies conducted within the framework of Diff Con theory (Svenson, 1992, 1996) have provided evidence for the existence of various forms of differentiation and consolidation processes, attempts to demonstrate that accountability determines the strength of such effects have been unsuccessful so far. As reported in Chapter 4, Svenson, Ortega Rayo, Andersen, Sandberg and Svahlin (1994) did not find any facts restructuring consolidation when participants were made accountable before making a decision. A possible reason for this failure to find the predicted stronger consolidation in this case may have been the more thorough information processing instigated by pre-decisional accountability, which may be assumed to have improved participants' memory for attribute outcomes and prevented facts restructuring from occurring. Although participants in the present study were urged to report their perception of the value of information at the time when they made their judgements, similar memory effects and/or the wish to appear consistent in their ratings cannot be ruled out and would have interfered with any differentiation and consolidation effects. Also, the prediction derived from Diff Con theory that accountable participants will bolster their chosen alternative more strongly than non-accountable participants may be too

simplistic, as it does not take into account any impression management concerns of accountable individuals, especially those who are externally accountable. The findings of the present study seem to consistently suggest that such concerns were important, especially to individuals who had been made externally accountable before the final decision. When the preferences of the external audience are salient and the decision maker is sufficiently motivated to show some regard for them, external accountability may indeed, as suggested earlier, result in a 'sitting on the fence' strategy, that is, a decreased projected evaluative difference between the alternative preferred by the decision maker and the alternative thought to be preferred by the external audience, in an attempt to appear less extreme. However, as discussed above, a different picture may emerge, when there is an opportunity for *hidden* bolstering of the own and devaluation of the audience's preferred alternative. It seems that individuals are willing to engage in such processes, provided they remain covert. Unfortunately, the present experiment, due to the absence of any significant effects, does not allow any firm conclusions in this matter, but the observed data pattern certainly points in this direction.

Justifications

Finally, analysis of participants' written justifications also revealed impression management strategies, which, as expected, were stronger, for participants who had been made externally accountable. Generally, participants presented consonant arguments much more often than they did dissonant arguments, with a ratio of almost 9:1 on average. The overall number of consonant arguments that were presented did not differ significantly between experimental conditions, but there was a significant interaction effect between time and type of accountability for a particular type of consonant argument, namely negative aspects of the non-chosen alternatives. When participants had been made accountable before the first decision, internally accountable participants mentioned relatively more such arguments than externally accountable participants. When participants had been made accountable before the final decision, on the other hand, the opposite tended to be true, although not significantly so. The number of arguments mentioning the chosen alternative in a positive way was not significantly affected by the experimental variables.

Externally accountable participants were found to generate relatively more arguments mentioning positive aspects of the non-chosen alternatives than internally accountable participants. Yet, where negative aspects of the chosen alternative were concerned (the other type of dissonant argument that could be produced), *internally* accountable participants did so significantly more often than externally accountable participants. These findings are consistent with the previously discussed idea of externally accountable participants being concerned about making a good impression; they avoided letting their own choice appear vulnerable, but showed that they were aware of the advantages of other alternatives preferred by their external audience. They may have also adopted a strategy of then refuting the alleged advantages of the non-chosen alternatives. Lending support to this assumption was the fact that participants who had been made externally accountable before the final decision tended to mention a relatively larger number of *negative* aspects of the non-chosen alternatives than participants under internal accountability. However, given that the number of arguments mentioning the non-chosen alternatives was generally very small, these results should be viewed with caution.

The other effect that could be observed was that participants under post-decisional rather than pre-decisional accountability generated a relatively smaller number of elaborations and of false reproductions supporting the chosen alternative. The most likely explanation for these effects is that these participants had not as much time to think about in depth and elaborate on as well as prepare strategic misrepresentations of information after having been made accountable, compared to participants who had already been made accountable before the first decision.

Finally, unlike expected, the justifications of participants who had been made accountable before the final decision were not significantly more complex than those of participants who had been made accountable before the first decision. Rather, a (non-significant) trend in the opposite direction emerged. Generally, the complexity of argumentation was very low, as participants did not engage in comparisons between alternatives very much, but mostly used one-sided arguments that only mentioned one alternative. This may have been partly due to the format of the justification, which asked participants to summarise their justification in keywords rather than give a verbal defence of their decision.

Conclusions

Although the experiment on the whole did not support the predictions of the PMA directly and consistently, it provided some valuable insights into the constraints that operate on the effects of internal and external accountability on decision processes, most notably the importance of individuals' attention to external norms and impression management concerns. As stated by the PMA, the implied presence of others and the knowledge of one's performance being identifiable, which is inherent in external accountability demands, creates evaluation apprehension and induces a motivation to make a decision that can be easily justified to the external audience and make a favourable impression on others. This will bias the decision maker's information processing towards supporting the alternative that is perceived to be most justifiable, or, when accountability is only introduced after the decision maker has already committed him- or herself to an alternative, towards defending this choice to the external audience. However, impression management concerns mean that the decision maker does not want to be seen doing this too openly. A similar suggestion has been made previously by authors like Kunda (1990), Baumeister and Newman (1994) and Hsee (1995, 1996). This fact has important implications for the measurement of information processing biases; they are most likely to be noticeable in covert measures, and not in measures where decision makers have to directly judge the quality of their preferred option compared to others. In the latter case, decision makers will keep the evaluative difference between their preferred alternative and other alternatives that may be preferred by their external audience small, in order to be able to change their choice in the face of external pressure without losing face should it be necessary, but also in order to signal that they take seriously any different opinion someone might have. This should be of particular importance to decision makers when there is normative ambiguity and decision makers cannot be sure about what reaction to expect from their audience, or when they anticipate dissent. This hypothesis can help to explain the failure of previous studies to find significant differentiation and consolidation effects as a consequence of accountability and provide a way forward of locating such effects in the future.

Because of the problem of covert differentiation and consolidation, the present study did not manage to highlight the differences between the effects of pre-decisional internal and external accountability sufficiently, as unidimensional values and weights were not elicited

after the first decision, and there was therefore no opportunity for externally accountable participants to engage in such processes. As to the difference between these two forms of accountability after an initial commitment had been made, the PMA predicted that it would be reduced, because both externally and internally accountable participants would have a directional goal of supporting their previously chosen alternative. The introduction of a stalemate between their preferred and another alternative, however, meant that external norms were made implicitly salient not only to externally but also to internally accountable participants. The present findings suggest, however, that for internally accountable participants this effect was not as strong as for externally accountable participants, and this meant that internally accountable participants tended to be more inward-looking and less balanced in their information processing if they had been made accountable before the final decision. This was evident in a number of information search measures, such as the lowest observed depth of search, highest observed variability of search across alternatives, lowest observed degree of compensatory search and strongest concentration of search on the chosen alternative. Internal accountability before the first decision, conversely, resulted in exactly the opposite effects, namely the highest observed depth of search, lowest observed variability of search across alternatives, highest observed degree of compensatory search and least strong concentration of search on the chosen alternative. Also, whereas participants who had been made internally accountable before the final decision showed the strongest bias in their final information search against the competing alternative, participants who had been made internally accountable before the *first* decision displayed the weakest such bias. This tendency for a balanced information processing for participants who had been made internally accountable before the first decision was also evident in some of the information evaluation measures. These participants, for example, displayed the smallest degree of differentiation between the chosen and the average non-chosen as well as competing alternative, both in actual and MAUT predicted ratings. Hence, in this particular situation, internal accountability before the first decision seems to have had the most beneficial effects. Despite the fact that participants had already been made accountable some time before their final decision, their accuracy goals still seemed to be strongly salient and their motivation to support their previously chosen alternative relatively low. Participants who had been made internally accountable before the *final* decision, on the other hand, would have had a stronger motivation to protect their initial choice and at the same time lack the motivation to prepare

for a public defence induced by external accountability, therefore having the freedom to show the strongest bias in their information processing.

Given the apparent significance of impression management and normative concerns, the second study, reported in the next chapter, focused on the effects of external accountability and adaptation to external audience preferences, depending on the extent of conflict between participants' own preferences and those of their audience. In a departure from the present study, external norms were made explicit and it was investigated how these norms influenced the decision maker's attention to and evaluation of particular attributes and alternatives. The paradigm used was similar to that adopted for the present study. In order to forego some of the problems encountered in the present study, however, some changes were made to the procedure. In order to allow participants to display covert differentiation and consolidation both after an initial and a final decision, unidimensional values and weights were elicited three times, before the first decision, after the first decision and after the final decision. These judgements were elicited not only for the attributes contained in the information set searched before the initial decision but also for their parallel attributes in the information set searched before the final decision, in order to see whether there would be any differentiation and consolidation for the latter as well. In addition, distracter attributes were introduced, in order to increase the number of judgements that had to be made to such an extent that it would be very difficult for participants to just remember their previous responses instead of reporting their preferences at the time when judgements were elicited. To eliminate the problem of any ceiling effects in the initial information search, the complexity of the task was reduced and participants were not told beforehand that they would have to judge each alternative as well. Also, the information presented in the second information set was highly redundant with that in the first, by expressing the information contained in the first set in a different format. Finally, in order to be able to compare the behaviour of participants who had been made accountable either before a first or before a final decision to that of non-accountable participants, a control group was introduced that did not receive any accountability instructions at any time.

CHAPTER 7: ACCOUNTABILITY TO REPRESENTATIVES OF DIFFERENT INTEREST GROUPS

Introduction

The aim of the following experiment was to study the norm-enforcement aspect of accountability, with respect to the mechanisms outlined in the PMA in Chapter 5. The model proposes that when the accountability manipulation suggests to individuals that their decision will be evaluated - by making them identifiable and providing certain norms or standards against which their performance can be measured - directional goals will become salient. As a consequence, individuals will bias their information evaluation and integration in such a way that the alternative that can be justified most easily will be favoured. Norms of evaluation may become salient through suggesting that the person or group to which individuals are accountable supports a particular view. For example, if individuals are told that they will have to justify their opinion or decision to a person with liberal views, they can anticipate that what they say will be evaluated in terms of how much it complies with liberal values. If the individual is made accountable to a conservative, on the other hand, the standard of evaluation is likely to be the extent to which expressed views are conservative. Tetlock's (1992) accountability model predicts that, in this case, individuals will employ an acceptability heuristic, whereby they adjust their view to that of their audience, that is, express liberal views when being made accountable to a liberal, and conservative views when being made accountable to a conservative (see Tetlock, Skitka, & Boettger, 1989). Similarly, the PMA assumes that individuals will evaluate information they have searched to help them make a decision in such a way that the alternative which would be supported by the person or group they are accountable to will come out best. This, however, will only be the case if they have not already got a strong view as to which alternative they prefer. If they have already committed themselves to a different alternative in a previous decision, individuals are assumed to bias the information evaluation process in such a way that their previously chosen alternative will fare best, rather than the alternative which would be preferred by the group

they are accountable to. This may be both for impression management reasons (individuals may not want to admit that their previous decision was wrong or appear opportunistic) and for cognitive economy reasons (it may be less laborious to create more reasons for why one's previous decision was the right one than to undo it and go through the whole decision process again). The tendency to adjust to the preferences of the external audience may be expected to be moderated by a number of factors, for example, the nature of the individual's own preferences and their strength, and the pressure perceived to be or actually exerted by the external audience. Also, as we have seen in the previous chapter, individuals are likely to avoid any open expression of bias when they experience normative ambiguity or anticipate pressure from the external audience.

Unlike in Tetlock's model, which predicts that individuals will not expend much cognitive effort when the views of the audience they are accountable to can be anticipated, the PMA proposes that feeling responsible for a decision will enhance the individual's wish to make an accurate decision and that information search will therefore become more extensive, especially with regard to information that may be regarded as important by the person or group the individual is accountable to. Again, this will only be the case if the individual has not already committed him- or herself to a decision. If a decision has already been made, the information search process is likely to become biased to support the previous decision.

In the experiment devised to test these hypotheses, the decision scenario presented to the participants came from the domain of Computer Science. The participants, who were all students of Computer Science, were asked to decide which of four CASE (Computer Aided Software Engineering) tools to buy for an insurance company. CASE tools are meta-programs, that is, programs that support the software development process, particularly the requirements analysis and design specification stage of this process.

In order to create a situation in which it was plausible for participants to be accountable to different interest groups, the following scenario was devised. Participants were asked to imagine being head of the information technology division of a company that wanted to acquire a CASE tool. The decision was said to be made by a committee which consisted of the participants themselves, a representative of the prospective users of the program, that is,

the company's programmers, and a representative of the company's management. Participants were told that, after having made a decision on their own, they would meet with the other committee members and discuss the alternatives in order to arrive at a joint decision. Participants were either made accountable to the users of the program or to the management of the company, by telling them that it would be particularly important to convince either the user representative (because the users would have to work with the program) or the management representative (because the management would have to make the necessary financial resources available) of their decision. This was expected to make either user benefits or cost concerns salient and constituted the 'type of accountability' manipulation.

The second factor manipulated was the time when the accountability manipulation was introduced, either before participants made a first decision or before they made a final decision. The two factors were fully crossed, resulting in four experimental groups. There was also a control group, members of which were not made accountable at any time.

Information search was measured by asking participants to search an information board on which information about four alternatives with regard to four attributes (two user attributes and two cost attributes) was presented, and recording their search process. Participants had to search information twice, once before they made a first decision and a second time before they made a final decision. The attributes and attribute outcomes for the second search were largely redundant with those of the first search; either the same fact was expressed slightly differently, for example, the user interface design was either described in words or expressed as a value on a satisfaction scale, or there was a natural correlation between the attributes, for example, the extent of training required and training costs. This was done so that participants would be able to anticipate the value of the information they searched, and so that it would be possible to test whether information search after a commitment was biased to support the previous conclusion. Information evaluation and integration was measured by participants' choices, their overall judgements of the suitability of each alternative and their judgements of attribute outcomes and attribute weights, which were elicited before participants made their first decision, after they had made their first decision and after they had made their final decision.

With regard to choices, it was predicted that participants who had been made accountable before the *first decision* would adjust their first decision to the group they had been made accountable to, that is, participants who had been made accountable to the users of the program would be more likely to choose alternatives with good outcomes on attributes which would be important to users of the CASE tool, whereas participants who had been made accountable to the management of the company would be more likely to choose alternatives with good outcomes on cost attributes. Participants who had not been made accountable were expected to make a decision that followed their own preferences, as indicated by their initial evaluation of unidimensional attribute outcomes and weight judgements. For the *final decision*, participants who had only been made accountable immediately before this decision were not expected to change their first decision and adjust to the preferences of the group they had been made accountable to, but were instead expected to stick to their initial decision. A choice in accordance with the preference of the group they had been made accountable to was therefore only expected for participants who had already been made accountable before the first decision, and for participants who had been made accountable before the final decision and whose own preferences coincided with those of the group they had been made accountable to.

A similar pattern was expected for participants' weight judgements. In line with Svenson's (1992, 1996) Diff Con theory, it was predicted that, after the *first decision*, participants who had been made accountable to the users of the program would report higher weights for user attributes than participants who had been made accountable to the management. The latter, on the other hand, would report higher cost attribute weights than participants who had been made accountable to the users. This importance differentiation would happen in order to adjust to the values of the group participants had been made accountable to and maximise the evaluative difference between their chosen and the other alternatives. In contrast, it was expected that non-accountable participants would only report high user attribute weights if their own preferences were user-oriented and high cost attribute weights if their own preferences were cost-oriented. After the *final decision*, participants who had been made accountable before the first decision were expected to further increase either their user attribute weights or their cost attribute weights, depending on which group they had

been made accountable to, whereas participants who had only been made accountable before the final decision would not show any adjustment to the preferences of the group they had been made accountable to. Instead, they were expected to increase the weights for attributes which supported their own preferences, especially if they were in a position of strong conflict, because their own preferences contradicted the preferences of the group they had been made accountable to.

Diff Con theory predicts that, as well as weights, participants will change their unidimensional value judgements in order to maximise the evaluative difference between their chosen alternative and their non-chosen alternatives. This tendency can be captured in the difference between the MAUT prediction for the chosen alternative and the average non-chosen alternative, since MAUT models combine unidimensional value and weight judgements when predicting the overall value of an alternative. Because this is a covert way of achieving differentiation and consolidation, since participants judge a multitude of single components that have to be combined by the experimenter in order to derive a model prediction, it was expected that, when given the opportunity, participants would engage in it. The MAUT predicted evaluative difference between the chosen and average non-chosen alternative was expected to be larger for participants who had been made accountable than participants who had not been made accountable, and larger for participants who experienced conflict because their own preferences contradicted those of the group they had been made accountable to than for participants who did not experience this conflict.

Conditions were also expected to differ in terms of the difference between MAUT predictions for alternatives complying with user preferences and MAUT predictions for alternatives complying with management preferences. The pattern of results was expected to follow the predictions for attribute weights. The differences should be larger for participants who had been made accountable to the users of the program than participants who had been made accountable to the management of the company, but only if they had been made accountable before the first decision. Non-accountable participants and participants who had been made accountable before the final decision were expected to show a larger difference if their own preferences were user-oriented rather than management-oriented, and particularly large differences if they were under additional pressure because they had been made

accountable to the management and therefore their own preferences clashed with those of the group they had been made accountable to.

Because participants' actual suitability judgements for the alternatives they had to choose between, including their ratings of user- and management-oriented alternatives, would make any differentiation and consolidation obvious to the external audience, they were not expected to show this bias *unless* they had complied with the preferences of the external audience. This was predicted to be more likely when participants' own attitude and that of their audience were the same. Hence, whereas the MAUT predictions were expected to show stronger evidence of differentiation and consolidation for participants who experienced a conflict between their own attitude and that of their external audience than participants who did not, the opposite was predicted to be evident in actual suitability ratings.

With regard to information search, it was expected that during both searches, accountable participants would search more information than participants who had not been made accountable. Accountable participants were also expected to show a lower variability of search across alternatives and attributes, to display more compensatory information processing in their search, to display a pattern of search that suggested less attribute-based processing and to show a lower concentration of search on the chosen alternative than non-accountable participants. For the final information search, those participants who had been made accountable only before this search were expected to display these tendencies to a stronger extent than participants who had already been made accountable before the first search.

In addition, if participants had been made accountable before the first search, those who had been made accountable to the users of the program were expected to search more user information, whereas participants who had been made accountable to the management of the company were expected to search more cost information. For non-accountable participants, on the other hand, the type of information they searched was expected to vary as a function of their own preferences. During the final information search, again only participants who had been made accountable before the first decision were expected to adjust the type of information they searched to the group they had been made accountable to.

Finally, it was expected that after a first decision had been made, information search would be biased to bolster the previously chosen alternative, that is, participants would search information which they could expect to support their chosen alternative and devalue their non-chosen alternatives. This tendency was expected to be stronger for accountable than for non-accountable participants and particularly strong if participants experienced a conflict because their own preferences opposed those of the group they had been made accountable to. Since it was expected that participants who had been made accountable before they made a commitment would have adjusted their preferences to those of the group they had been made accountable to, their information search was expected to be less biased than that of participants who had only been made accountable *after* they had made a previous commitment, especially if their own attitude coincided with that of the group they had been made accountable to.

Method

Participants

Participants were 80 students of Computer Science at the Technical University of Darmstadt, 73 males and 7 females. The number of semesters they had studied Computer Science ranged from 1 to 21 semesters, with a mean of 7.79. For 66 participants, Computer Science was their main subject, for 14 participants it was a minor subject. The participants' age ranged from 20 to 35 years, the mean age was 24.28 years. Participants received DM 10 as remuneration for their participation, or a lottery ticket with which they could win between DM 20 and DM 500. After completion of the data collection, 13 winners were drawn.

Design

Two factors were manipulated, type of accountability (to the company's users of a software program vs. the management of the company) and time of introducing accountability (before a first decision vs. after a first and before a final decision). Both factors were varied

between subjects, resulting in four experimental groups with 16 participants per group. A control group with 16 participants who were not made accountable at any time was also employed. In addition, participants' attitude (pro-user vs. pro-management) was included as a between-subjects factor in the analyses.

Materials

The decision problem put to the participants involved selecting one out of four CASE (Computer Aided Software Engineering) tools. The description of the decision scenario, attributes and attribute outcomes was developed with the help of two experts, one information technology consultant and one member of staff at the Computer Science department at the Technical University of Darmstadt. The decision scenario was closely modelled on a typical decision problem encountered by the consultant in her daily practice.

Like in the previous experiment, in order to be able to assess the extent of bias in information search after a previous commitment, it was necessary to construct two sets of redundant information. These sets had to fulfil certain criteria. Firstly, the information contained in the second set had to be redundant with the information contained in the first set, in such a way that participants would be able to predict the value of the outcome of a particular alternative on a particular attribute from their knowledge of the alternative's outcome on the parallel attribute. Secondly, in order to be able to test whether participants would utilise different information, depending on the group they had been made accountable to, the information sets had to contain an equal number of attributes that would be regarded as important either by the users of a CASE tool or by the management of the company buying the tool. Thirdly, in order to control for effects of format of description of attribute values (Huber, 1980; Stone & Schkade, 1991), attributes had to meet the condition that their outcomes could be expressed both in a verbal and in a numerical format[1].

[1] Stone & Schkade (1991), in an information board study, found that verbal descriptions of attribute outcomes resulted in a less attribute-based and more alternative-based information search, and less compensatory processing compared to numerical descriptions. This can be explained by comparisons within attributes being more difficult and effortful for words than for numbers. This is supported by a study by Huber (1980), who found

Four alternatives, defined by their outcomes on two sets of four parallel attributes, were constructed. Two of the four attributes in the first set were user attributes (user interface design and method support), the other two attributes were cost attributes (hardware costs and training requirements). The parallel attributes in the second set were satisfaction with the user interface design, satisfaction with the method support, hardware requirements, and training costs, respectively. Each alternative had a positive outcome on two attributes and a negative outcome on the two other attributes (see Table 7.1)[2]. This was done in order to make the overall attractiveness of the different alternatives as similar as possible to each other if an equal weighting of attributes was applied, while at the same time creating alternatives that would be differently attractive to participants if they reacted to being made accountable to a particular group by weighting the attributes assumed to be important to that group more heavily than the other attributes.

Table 7.1. The valence of attribute outcomes for each alternative in information set 1*.

Alternative	A	B	C	D
user attribute 1 (verbal description)	-	+	-	+
user attribute 2 (numerical description)	+	+	-	-
cost attribute 1 (verbal description)	-	-	+	+
cost attribute 2 (numerical description)	+	-	+	-

Note: * This was the same for information set 2, except that attribute outcomes expressed in a numerical format in information set 1 were now expressed in verbal terms and vice versa.
+ stands for a positive outcome, - for a negative outcome.

For reasons of counterbalancing, information set 1 contained information about the user interface design (verbal), satisfaction with the method support (numerical), training requirements (verbal), and hardware costs (numerical), while information set 2 contained

that in participants' verbal protocols, direct comparisons were more frequent with numerical data, while evaluative statements were more frequent with verbal descriptions.

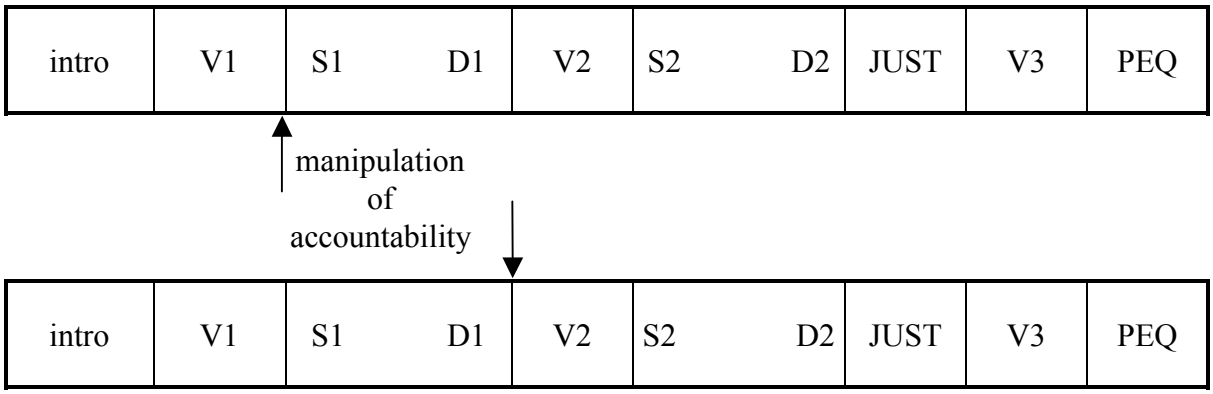
[2] Negativity in this context refers to a less positive outcome. To support the cover story, which stated that the alternatives to be considered by the participants had been pre-selected from a larger pool of alternatives, none of them had clearly negative features.

information about the satisfaction with the user interface design (numerical), method support (verbal), training costs (numerical) and hardware requirements (verbal) (see Appendices 4 and 5 for a full description of the alternatives).

Procedure

Table 7.2 shows a representation of the sequence of events during the experiment. They are explained in more detail below.

Table 7.2. Procedure.



Note: Capital letters indicate elicitation parts, lower-case letters indicate instruction parts.
intro: introduction to the decision problem
V1, V2, V3: elicitation of unidimensional values and weights
S1, S2: information search
D1, D2: decision
JUST: justification
PEQ: post-experimental questionnaire

Participants took part in the experiment in individual sessions. Written instructions and response sheets were handed out to them by the experimenter, who was present throughout the experiment and gave additional explanations if necessary. Participants were first introduced to the decision problem:

Intro: Introduction to the decision problem

Participants were asked to imagine being head of the information technology division of a large German insurance company. They were told that the company had suffered economic losses through stiff competition from foreign insurance companies. As a result, internal organisational procedures were to be revised to make them more efficient; this revision required new software developments. In order to support software development, it had been decided to acquire a CASE tool. Following some additional information about the

current programming environment and the software development method the department planned to use, participants were told that four tools had been pre-selected by a team of testers. The decision about which CASE tool to buy would be made by a committee consisting of themselves, a representative of the prospective users of the program, that is, the company's programmers, and a representative of the company's management. Given that the other two committee members represented different interests, it was expected that participants would feel that their own preference would be decisive and therefore feel responsible for the decision they made.

First, participants' preferences, that is, unidimensional values and weights, were elicited for 12 attributes, 6 user and 6 cost attributes. Two attributes in each set, that is, 4 of the 12 attributes, were distracters, two of them user-oriented (vendor support and documentation of results) and two management-oriented (tool price and vendor's reputation). They were not used in the later description of alternatives.

V1: First elicitation of unidimensional values and weights

After going through an example that illustrated how value judgements had to be made, in order to familiarise participants with the procedure, four outcomes were presented for each of the 12 attributes, which had to be rated on a scale from 0 (very bad) to 100 (very good). Participants were not aware of the fact that some of these were the outcomes which would later be used to describe the alternatives. The order of presentation of attributes as well as the order of presentation of attribute outcomes within each attribute was randomised for each participant.

Weights for the 12 attributes were elicited with the swing method developed by von Winterfeldt and Edwards (1986, see Chapter 6 for a detailed description). To avoid making the redundancy of some of the attributes salient to participants, weights were elicited in two separate sets, one consisting of 6 and the other of 7 attributes. Each set contained the attributes of one of the information sets plus two or three distracters. One distracter attribute included in the first weight elicitation set, namely 'vendor support', was added to the attributes of the second set, in order to be able to relate the weights obtained in the first set to the weights obtained in the second set. The order of presentation of attributes within each weight elicitation set was randomised for each participant.

Following the elicitation of unidimensional values and weights, accountability was manipulated in the two groups for which accountability was to be introduced before a first decision would be made.

Manipulations of accountability

Participants were reminded that they had been asked to assume the role of head of the information technology division, and that the final decision about which CASE tool to buy would be made by the project committee, consisting of themselves and two other members. They were told that each committee member was meant to first arrive at a decision on their own, before the committee members would meet, discuss the

alternatives and agree on which CASE tool to acquire for the project. It was made explicit that, after the first part of the experiment, participants would meet with two other participants who had been asked to assume the roles of the other committee members and that, following a group discussion, they were expected to arrive at a joint decision.

Accountability to users. Participants in these conditions were told that in the later group discussion, it would be particularly important to convince the representative of the users of the correctness of their own decision, since the users would be the people who had to work with the tool.

Accountability to management. Participants in these conditions were told that in the later group discussion, it would be particularly important to convince the representative of the management of the correctness of their decision, since the management would have to make the necessary financial resources available.

In order to make the accountability manipulation more credible and more powerful, participants were asked to sign a form declaring their consent that the group discussion was tape-recorded[3]. The tape-recorder was visible all the time, lying on the table at which the participants filled in their response sheets, at a distance of ca. 50 cm.

No accountability. Participants in the control group and in the two conditions in which accountability was only manipulated after a first decision had been made received none of the above information and they were not asked to sign a consent form. Also, no tape-recorder was visible to them.

After the accountability manipulation, participants were provided with instructions about how to search information about the four available alternatives and asked to make a decision for one of them.

S1: Information search set1

In order to monitor information search, the information board technique (Payne, 1976; see Chapter 6 for a more detailed description) was used. Information about four tools, named A, B, C, and D, was presented in an attribute-by-alternative matrix, with the alternatives constituting the columns and the attributes constituting the rows of the matrix. The information items were printed on cards and initially not visible. Participants had to turn over a card in order to be able to see the information given on it, and put it back again so that the information was not visible anymore before turning over the next card. Participants were free to search as many items of information as they wanted, and were allowed to look at the same item as often as they liked. The sequence in which cards were turned over was recorded by the experimenter.

[3] This is a manipulation that has first been suggested and employed by Tetlock (1983a, b).

D1: First decision and suitability judgements

Participants indicated which tool they would choose and also rated each tool in terms of how suitable it was for the company. This was done on a scale ranging from 0 (extremely unsuitable) to 100 (extremely suitable). Participants had not been told in advance about these ratings, in order to avoid an effect on the decision process, which has been shown to be more compensatory when ratings of alternatives are called for compared to making a choice (see Westenberg, 1991, for a review). In addition, participants rated how difficult the decision had been (on a scale ranging from 1 'very easy' to 7 'very difficult'), how confident they were about having made a good decision (on a scale ranging from 1 'very unsure' to 7 'very confident') and how strong their need had been to make a good decision (on a scale ranging from 1 'very weak' to 7 'very strong').

After participants had made their decision and provided their ratings, unidimensional values and weights were elicited for a second time.

V2: Second elicitation of unidimensional values and weights

Participants rated the outcomes of all attributes including the distracter attributes as they had done at the beginning of the experiment. Weights were also elicited in the same way as before.

This was followed by the manipulation of accountability for participants in the 'accountability before the final decision' conditions.

Manipulation of accountability

Accountability either to the users of the tool or the management of the company was manipulated in the same way as described above. Participants in the control group and in the 'accountability before the first decision' conditions did not receive any information at this point.

Subsequently, participants were given the opportunity to search additional information about the tools and told to make a final decision.

S2: Information search set2

As before, the information that could be searched was presented in form of a matrix on an information board. In the upper half of the board the information items participants had looked at during their first search were still present and permanently visible; the information items they had not searched were removed from the board, so that they were not accessible anymore. Participants searched information in the same way as before; the sequence of search was again recorded by the experimenter.

D2: Final decision and suitability judgements

Participants announced their final decision, rated the suitability of each alternative again and responded to questions requiring them to rate the difficulty of the decision, their certainty about having made a good decision and the strength of their need to make a good decision.

This was followed by a written justification of their choice.

JUST: Justification

Participants had to justify their choice in writing. Participants in the control group were asked to state the reasons for their choice, participants in the experimental groups were asked to write down the reasons they would give in their discussion with the other committee members.

The justification was followed by the third and last elicitation of unidimensional values and weights.

V3: Third elicitation of unidimensional values and weights

The third elicitation of unidimensional values and weights followed the procedure described above.

PEQ: Post-experimental questionnaire

At the end of the experiment, participants were asked to provide a number of ratings, which partly served as manipulation checks. First, they rated the similarity between pairs of attributes on a 5-point scale, ranging from 1 ‘very dissimilar’ to 5 ‘very similar’. The pairs consisted of all possible combinations of the 8 attributes contained in information sets 1 and 2, altogether 28 of them. The order of presentation of the 28 pairs of attributes was randomised for each participant.

Finally, participants rated

- how much they had expected to have to justify their decision (on a scale from 1 ‘not at all’ to 7 ‘very much’)
- how strongly they had felt a need to be able to justify their first decision (on a scale from 1 ‘very weak’ to 7 ‘very strong’)
- how strongly they had felt a need to be able to justify their final decision (on a scale from 1 ‘very weak’ to 7 ‘very strong’)
- how useful they had found the information that could be accessed during the final information search (on a scale from 1 ‘not all useful’ to 7 ‘very useful’).

Results

Similarity between attributes

Analysis of the similarity ratings participants provided at the end of the experiment showed that the attempt to create information sets with redundant attributes had been successful. The mean similarity ratings of pairs of redundant attributes all exceeded a value of 4, suggesting that these attributes had been perceived by the participants as highly similar on average. None of the other attribute pairings showed mean ratings that came close to those for the parallel attributes (see Appendix 3 for a complete list of results). Table 7.3 shows mean normalised weights and similarity ratings for the attributes that were used.

Table 7.3. Mean normalised weights and similarity ratings for the attributes used in both information sets.

Information Set 1			Information Set 2	
Attribute Labels	*Weight*	*Similarity*	*Weight*	Attribute Labels
user interface design	0.37	4.82	0.33	satisfaction with user interface design
hardware costs	0.12	4.76	0.15	hardware requirements
satisfaction with method support	0.33	4.70	0.34	method support
training requirements	0.17	4.35	0.17	training costs

Note: Similarity ratings were given on a scale from 1 (very dissimilar) to 5 (very similar).

Attitude

Since many of the following analyses include participants' attitude as an explanatory variable, it will first be explained how attitude was measured and on what basis participants were classified as having either an attitude that supported user values (a pro-user attitude) or an attitude that supported management values (a pro-management attitude). The attitude measure was derived from participants' initial unidimensional value and weight judgements, which had been elicited before any experimental manipulation. More specifically, for each participant the MAUT prediction of the overall value of the alternative with exclusively positive outcomes on the user attributes and exclusively negative outcomes on the cost attributes (alternative B, pure user alternative) was calculated. This prediction combined the normalised weight and value judgements for the eight experimental attributes and their respective outcomes by feeding them into the simple linear additive MAUT model specified in equation 1 (see Chapter 1). It was assumed that someone whose attitude was inclined towards user considerations would give preference judgements which resulted in a high overall predicted value for the pure user alternative, whereas the preference judgements of someone who was more management-oriented should have resulted in a low predicted overall value for this alternative. Therefore, a median split was carried out on the MAUT prediction for the pure user alternative, in order to classify participants according to their attitude. The obtained median was high (66.32), indicating that participants generally attached high weight to user attributes[4] . For participants with a value below the median, however, this tendency was less marked, and they will therefore be referred to as having a pro-management attitude. Participants with a value above the median will be referred to as having a pro-user attitude.

[4] This was not surprising, given that in their training as programmers user considerations featured highly.

First decision

Information search

The examined process variables were the same as described in experiment 1. The effects of accountability and attitude on these variables and on the type of searched information were analysed and are summarised in Table 7.4.

1. Depth of search

Information search was generally exhaustive. The overall mean percentage of total number of information items searched was 115 %. This means that some information items must have been accessed at least twice. The overall mean percentage of different cells searched was 83%, on average 13.35 of the 16 different available items were accessed.

It had been predicted that accountable participants would show a deeper information search, that is, search more information than non-accountable participants. To test this hypothesis, the 'accountability to users' and 'accountability to the management' conditions were pooled and compared to the 'no accountability' conditions. A two-factorial ANOVA, testing the effects of accountability and attitude, was performed on both the percentage of total number of items searched and the percentage of different cells searched. The expected accountability main effect was not significant for either of these measures, although the observed differences between conditions were in the predicted direction (total number of items searched: accountability: $M = 123\%$, no accountability: $M = 111\%$; different cells searched: accountability: $M = 84\%$, no accountability: $M = 82\%$). In both cases, there were significant main effects of attitude, however (total number of items searched: $F(1,76) = 8.26$, $p = .005$; different cells searched: $F(1,76) = 8.12$, $p = .006$). These effects will be discussed in more detail below, as part of the analyses of the effects of different types of accountability.

Table 7.4. Process measures for the first information search, as a function of type of accountability and attitude.

	Accountability					
	No		To Users		To the Management	
	Pro User (n=24)	Pro Man. (n=24)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)
1. Depth of search (in %)						
total # of items	101	122	108	107	90	167
different cells	77	88	88	82	68	95
2. Searched user information (in %)						
total # of items[1]	62	55	63	62	64	54
different cells[2]	60	53	58	58	60	51
3. Variability of search across alternatives						
total # of items	1.41	1.39	1.16	1.33	1.64	1.75
different cells	0.78	0.64	0.49	0.59	0.98	0.28
across attributes						
total # of items	1.29	1.29	1.52	1.10	1.37	1.23
different cells	0.69	0.39	0.62	0.55	0.84	0.20
4. Compensatory processing	0.50	0.63	0.68	0.62	0.37	0.83
5. Pattern of search (Payne Index)	-0.49	-0.22	-0.59	-0.19	-0.26	-0.22
6. Concentration of search						
total # of items	0.41	0.24	0.32	0.42	0.58	0.27
different cells	0.28	0.17	0.16	0.18	0.50	0.06

[1] For each participant, the number of user items searched was divided by the total number of items they had searched and transformed into a percentage.
[2] For each participant, the number of user information cells accessed was divided by the number of different cells they had searched and transformed into a percentage.

In order to test the effects of type of accountability and attitude, 3 (type of accountability: no vs. to users vs. to the management) by 2 (attitude: pro-user vs. pro-management) ANOVAs were conducted. The analyses showed a significant main effect of participants' attitude on the percentage of total number of information items searched ($F(1,74)$

= 7.66, p = .007). Participants with a pro-management attitude searched a higher percentage of information than participants with a pro-user attitude (M = 130% vs. M = 101%). The reason for this initially unexpected result could be that, in contrast to participants with a pro-user attitude, who only focus on user attributes and therefore search relatively little information, participants with a pro-management attitude cannot afford to only focus on cost attributes. Instead, cost-benefit trade-offs have to be made, in order to identify the product that is reasonably good and not too expensive at the same time. This would mean that people with a pro-management attitude have to take more factors into account than people with a pro-user attitude and therefore should search more information. If this was the case, participants with a pro-management attitude should have also searched more different cells than participants with a pro-user attitude. This was indeed the case (M = 88% vs. M = 78%, F(1,74) = 7.37, p = .008). Contrary to the original predictions, there was no main effect of type of accountability on the percentage of total number of items searched, but a significant interaction between type of accountability and attitude was revealed (F(2,74) = 3.18, p = .047, see Figure 7.1). A simple main effects analysis showed that the accountability manipulation did not significantly affect participants with a pro-user attitude, but did have a significant effect on participants with a pro-management attitude (F(2,74) = 4.19, p = .019). The latter searched a higher percentage of information if they had been made accountable to the management of the company than if they had either not been made accountable at all (M = 167% vs. M = 122%, t(79) = 2.51, p = .014) or had been made accountable to the users of the program (M = 167% vs. M = 107%, t(79) = 2.61, p = .011). Also, the effect of attitude depended on the accountability condition participants were in (F(3,74) = 4.61, p = .005). The only significant difference was found in the 'accountability to management' condition, where participants with a pro-management attitude searched a higher percentage of information than participants with a pro-user attitude (M = 167% vs. M = 90%, t(79) = 3.34, p = .001).

A significant interaction between type of accountability and attitude was also obtained for the percentage of different cells searched (F(2, 74) = 4.30, p = .017). A simple main effects analysis indicated that participants with a pro-user attitude were significantly affected by the accountability manipulation, whereas participants with a pro-management attitude were not (F(2,74) = 3.20, p = .047). Participants with a pro-user attitude searched a higher

percentage of different cells when they had been made accountable to users, compared to when they had been made accountable to the management (M = 88% vs. M = 68%, $t(79)$ = 2.50, p = .015). In addition, the effect of attitude was dependent on the accountability manipulation ($F(3,74)$ = 6.03, p = .001). Participants in the 'no accountability' and 'accountability to the management' conditions searched a higher percentage of different cells when they had a pro-management attitude than when they had a pro-user attitude (no accountability: M = 88% vs. M = 77%, $t(79)$ = 2.37, p = .021; accountability to management: M = 89% vs. M = 68%, $t(79)$ = 3.47, p = .001). There was no significant difference between participants with a pro-management and a pro-user attitude in the 'accountability to users' conditions. Again, the expected main effect of type of accountability on percentage of different cells searched was not significant.

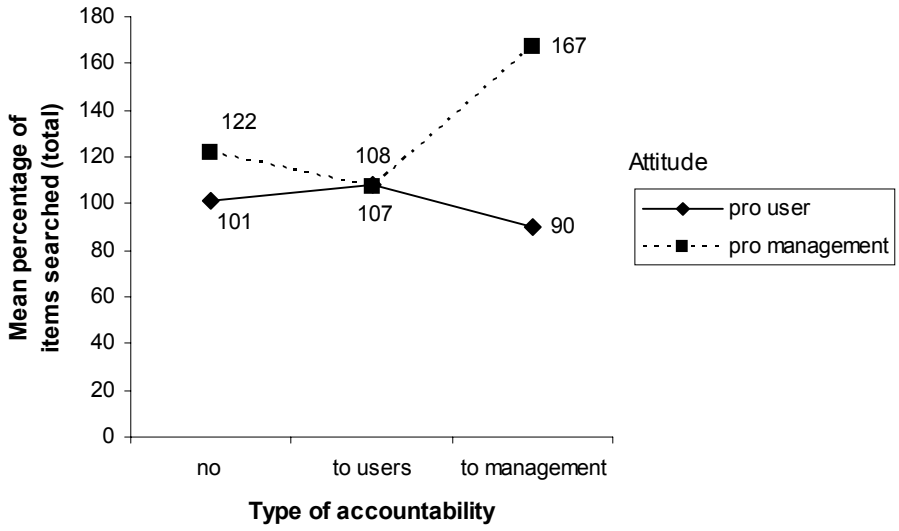


Figure 7.1. Mean percentage of items searched during the first information search (based on the total number of items searched), as a function of type of accountability and attitude.

2. Searched user information

The items searched were further classified as to whether they related to user information or whether they related to cost information. The proportion of information searched that was related to user attributes (user interface design and satisfaction with method support) and to cost attributes (hardware costs and training requirements), out of the total number of information items searched and out of the number of different cells searched, was calculated and transformed into a percentage. A balanced search of both types of information would have resulted in a 50:50 split. Only the results relating to user information are presented in Table 7.4, since the results for cost information represent the complimentary values to arrive at 100% and are therefore redundant.

Overall, participants preferred user information over cost information, although this difference was less pronounced for the different cells searched than for the total number of items searched (*M*(user/total) = 60%, *M*(cost/total)=40%, *M*(user/different cells) = 56%, *M*(cost/different cells) = 44%). The fact that user information was preferred is consistent with the fact that, generally, participants' attitudes were more user-oriented.

It had been expected that participants who had been made accountable to the users of the program would search a higher percentage of user information than participants who had not been made accountable or who had been made accountable to the management, resulting in a main effect of type of accountability. Two-factorial ANOVAs testing the effects of type of accountability and attitude were performed on the percentage of user information related to the total number of items searched as well as different cells searched. The only significant effects obtained were significant main effects of attitude on the percentage of searched user information for the total number of items searched ($F(1,74) = 5.35$, $p = .024$) and the percentage of searched user information for different cells searched ($F(1,74) = 4.67$, $p = .034$). As expected, participants with a pro-user attitude searched more user information than participants with a pro-management attitude (total number of items searched: $M = 63\%$ vs. $M = 56\%$, different cells searched: $M = 59\%$ vs. $M = 53\%$). There were neither significant main effects of type of accountability nor significant interactions between type of accountability and attitude on both measures, although the obtained pattern of results was consistent with

expectations. For the percentage of searched user information related to the total number of items searched as well as different cells searched, simple main effect analyses indeed indicated that the effect of attitude was dependent on the particular accountability condition (total number of items searched: $F(3,74) = 3.04$, $p = .034$, different cells searched: $F(3,74) = 3.54$, $p = .019$). A significant and marginally significant difference, respectively, between participants with a pro-user and a pro-management attitude was only found for participants in the 'no accountability' and the 'accountability to the management' conditions. In these conditions, participants with a pro-user attitude searched a higher percentage of user information than participants with a pro-management attitude (total number of items searched: no accountability: $M = 62\%$ vs. $M = 55\%$, $t(79) = 2.37$, $p = .020$, accountability to the management: $M = 64\%$ vs. $M = 54\%$, $t(79) = 1.84$, $p = .070$; different cells searched: no accountability: $M = 60\%$ vs. $M = 53\%$, $t(79) = 2.63$, $p = .010$, accountability to the management: $M = 60\%$ vs. $M = 51\%$, $t(79) = 1.92$, $p = .059$). Also, as expected, the percentage of searched user information for participants with a pro-management attitude tended to be higher when they had been made accountable to the users of the program, compared to when they had not been made accountable (total number of items searched: $M = 62\%$ vs. $M = 55\%$, $t(79) = 1.44$, $p = .155$, different cells searched: $M = 58\%$ vs. $M = 53\%$, $t(79) = 1.44$, $p = .153$) or had been made accountable to the management (total number of items searched: $M = 62\%$ vs. $M = 54\%$, $t(79) = 1.48$, $p = .143$, different cells searched: 58% vs. 51%, $t(79) = 1.54$, $p = .127$). The mean percentage of searched user information (based on the total number of items searched) observed in the different conditions is shown in Figure 7.2.

3. Variability of search

The overall mean variability of search across alternatives and attributes, computed for the search of different cells, was 0.65 across alternatives and 0.54 across attributes. The fact that both values are greater than 0 indicates that participants did not search an equal number of attributes for each alternative and did not search an equal number of alternatives with respect to each attribute. Variability of search computed for the total number of items searched takes repeated searches of the same cells into account and therefore can shed more

light on the kind of information that received particular attention when information was searched more than once. The overall mean variability of search across alternatives computed for the total number of items searched was slightly higher than the overall mean variability of search across attributes (M = 1.43 vs. M = 1.30). This indicates that alternatives were attended to more unevenly than attributes, when the same information was searched more than once.

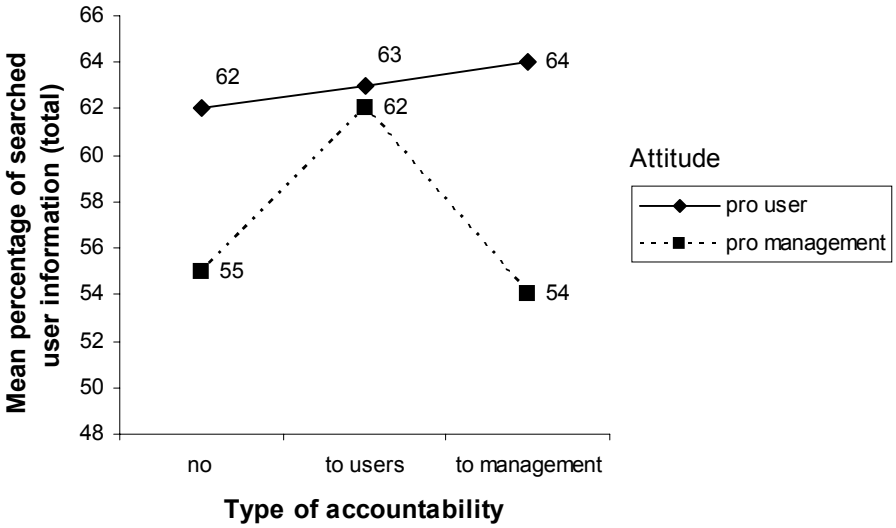


Figure 7.2. Mean percentage of searched user information (based on the total number of items searched), as a function of type of accountability and attitude.

It had been predicted that accountable participants would show a lower variability of search for both alternatives and attributes than participants who had not been made accountable. This tendency was emergent in the data only for variability of search across alternatives for different cells searched (accountability: M = 0.56, no accountability: M = 0.71), but did not reach standard significance levels. The variability of search across alternatives computed for the total number of items searched was, in fact, higher for accountable participants than non-accountable participants, but, again, not significantly so (M = 1.47 vs. M = 1.40). The variability of search across attributes, both for the total number of items searched and for different cells searched, was virtually the same for accountable and non-accountable participants (total number of items searched: M = 1.31 vs. M = 1.29, different cells searched: M = 0.54 vs. M = 0.54). The only significant effect in the analyses, a

main effect of attitude, was obtained for variability of search across attributes computed for different cells searched ($F(1,76) = 10.01$, $p = .002$)[5]. It will be discussed in more detail below.

The ANOVAs testing the effects of type of accountability and attitude on variability of search across alternatives, computed for the total number of items searched as well as different cells searched, showed no significant effects. The same was true for variability of search across attributes computed for the total number of items searched. There was, however, a significant main effect of attitude on the variability of search across attributes for different cells searched ($F(2,74) = 8.26$, $p = .005$). Participants with a pro-user attitude showed a significantly higher variability of search across attributes ($M = 0.70$) than participants with a pro-management attitude ($M = 0.38$). A high variability of search across attributes is obtained if the number of cells (i.e., alternatives) inspected for each attribute is very different and is an indication of non-compensatory processing. The high variability of search across attributes for participants with a pro-user attitude therefore suggests that they were more likely to adopt a non-compensatory decision strategy, where they stopped to consider an alternative once they discovered that it did not meet certain criteria. Participants with a pro-management attitude, on the other hand, seem to have considered a larger number of alternatives before narrowing their options down, and, as a result, searched the information set more extensively (as supported by the results for depth of search).

4. Compensatory processing

The overall mean Koele and Westenberg Index of Compensatory Processing (see Chapter 6) was 0.60, where a maximum extent of compensatory processing is indicated by a value of 1. This suggests that, overall, information processing was non-compensatory to some extent.

It had been expected that accountable participants would process information in a more compensatory manner than participants who were not accountable. Although the

[5] A weak tendency for an attitude main effect was observed for variability of search across alternatives, computed for different cells searched ($F(1,76) = 2.74$, $p = .102$, pro-user: $M = 0.75$, pro-management: $M = 0.55$)

observed means did suggest a higher degree of compensatory information processing for accountable participants compared to non-accountable participants (accountability: $M = 0.64$, no accountability: $M = 0.57$), this difference was not statistically reliable, as there was no significant main effect of accountability in the ANOVA conducted to test this hypothesis. There was a significant main effect of attitude, however ($F(1,76) = 4.47$, $p = .038$). Participants with a pro-management attitude processed information in a more compensatory manner than participants with a pro-user attitude ($M = 0.67$ vs. $M = 0.52$). This result was not surprising, given that the index combines depth of search and variability of search across alternatives for different cells, and that participants with a pro-management attitude had been found to search more information and show a tendency towards a lower variability of search across alternatives for different cells than participants with a pro-user attitude.

The analysis testing the effects of type of accountability and attitude on the Index of Compensatory Processing revealed the already mentioned significant main effect of attitude ($F(1,74) = 4.49$, $p = .037$), but no other significant effect.

5. Pattern of search (Payne Index)

The mean Payne Index across all conditions was -0.34, suggesting that, overall, information search was attribute-wise rather than alternative-wise, that is, participants tended to search cells for the same attribute before moving on to the next one rather than searching different attributes for the same alternative before moving on to the next alternative. The difference to the expected value of 0 for a balanced search was highly significant ($t(79) = 6.25$, $p = .000$).

It had been expected that the pattern of search would be more attribute-wise for non-accountable than for accountable participants. This is because non-accountable participants were expected to have a lower motivation to be accurate in their processing than accountable participants, and would therefore be more likely to resort to this cognitively easier strategy, which is commonly associated with non-compensatory processing (Russo & Dosher, 1993). Participants who were not accountable did indeed show a slightly stronger attribute-wise search than participants who were accountable ($M = -0.36$ vs. $M = -0.33$), but this difference

was not significant. There was a significant main effect of attitude on the pattern of search, however ($F(1,76) = 5.43$, $p = .022$). Participants with a pro-user attitude searched information in a more attribute-wise fashion than participants with a pro-management attitude ($M = -0.47$ vs. $M = -0.21$). This further supports the conclusions drawn from the results obtained for variability of search across attributes and the Index of Compensatory Processing, which suggested that participants with a pro-user attitude processed information in a more non-compensatory manner than participants with a pro-management attitude.

Type of accountability did not affect the Payne Index significantly. In the ANOVA testing its effect, which also included the attitude factor, the main effect of attitude was only marginally significant ($F(1,74) = 3.68$, $p = .059$).

6. Concentration of search

The fact that the overall amount of information searched was different for different alternatives (which had already been indicated by the results for variability of search across alternatives and the Index of Compensatory Processing) was further supported by the concentration of search measure. Mean concentration of search on the chosen alternative across all conditions was 0.35 for the total number of information items searched and 0.22 for different cells searched. The fact that these values are not 0 shows that the chosen alternative generally received more attention than the non-chosen alternatives.

It had been predicted that non-accountable participants would concentrate their search more on the chosen alternative than accountable participants. Such a trend could neither be observed for the total number of items searched (no accountability: $M = 0.33$, accountability: $M = 0.39$) nor for different cells searched (no accountability: $M = 0.22$, accountability: $M = 0.21$). Rather, the observed pattern of results suggests that when information was searched repeatedly, accountable participants were more biased towards their chosen alternative than non-accountable participants. However, in neither case was the main effect of accountability significant. There were significant main effects of attitude on both measures, however (total number of items searched: $F(1,76) = 4.30$, $p = .041$; different cells searched: $F(1,76) = 7.36$, $p = .008$), the nature of which will be discussed in more detail below.

The analyses testing the effects of type of accountability and attitude on the different concentration of search measures identified a significant main effect of attitude on concentration of search for different cells searched ($F(1,74) = 8.47$, $p = .005$). Consistent with the previously reported results, participants with a pro-management attitude concentrated their search less on the chosen alternative than participants with a pro-user attitude ($M = 0.15$ vs. $M = 0.29$)[6]. In addition to the attitude main effect, a significant interaction between attitude and type of accountability was observed ($F(2,74) = 4.34$, $p = .016$). This is illustrated in Figure 7.3. Simple main effects analyses indicated that the accountability manipulation only affected participants with a pro-user attitude. They showed a stronger concentration of search on the chosen alternative when they were accountable to the management than when they were accountable to the users of the program ($M = 0.50$ vs. $M = 0.16$, $t(79) = 2.89$, $p = .005$). Also, the effect of attitude was found to depend on the accountability condition participants belonged to ($F(3,74) = 5.57$, $p = .002$). A significant difference between participants with a pro-user and a pro-management attitude was obtained in the 'accountability to management' condition only. Participants with a pro-user attitude concentrated their information search more on the eventually chosen alternative than participants with a pro-management attitude ($M = 0.50$ vs. $M = 0.06$, $t(79) = 3.74$, $p = .000$). This may be explained by the fact that the former were in a situation of high conflict, because their attitude was different from that of the group they had been made accountable to, and seem to have been reluctant to solve the conflict by adjusting to the latter's view. Participants with a pro-management attitude who had been made accountable to the users of the program, in contrast, who also experienced conflict between their own attitude and that of the group they had been made accountable to, did not show an equivalent effect (although their concentration of search was also higher than for participants with a pro-user attitude, who did not experience conflict). This suggests that they experienced the conflict between their own preferences and those of their audience less strongly.

[6] A similar tendency was observed for concentration of search for the total number of items searched (pro-management: $M = 0.28$, pro-user: $M = 0.42$, but the effect was only marginally significant in this case ($F(1,74) = 3.16$, $p = .080$).

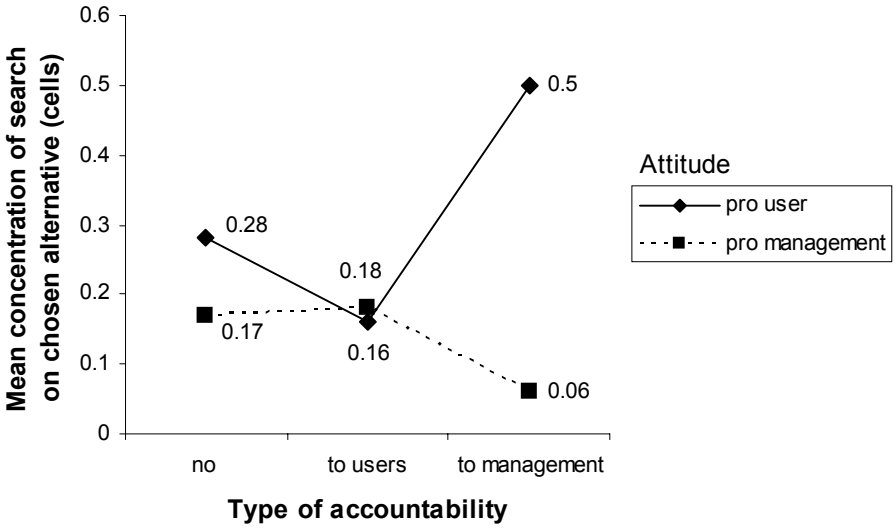


Figure 7.3. Mean concentration of the first information search on the chosen alternative (based on the different cells searched), as a function of type of accountability and attitude.

Summary of the results for the first information search

The pattern of results for the first information search suggests that participants did react to the accountability manipulation, but only significantly so, if they had a pro-management attitude. In this case, they searched more information when made accountable to the management and adapted their information search by searching less information when made accountable to the users of the program. The results obtained for the type of information searched are consistent with this interpretation; when made accountable to the users of the program, participants with a pro-management attitude tended to search relatively more user information, when made accountable to the management of the company they tended to search relatively less user information than under no accountability. In the 'no accountability' condition, information search followed participants' attitude; when participants had a pro-user attitude they searched less information than when they had a pro-management attitude. When participants with a pro-user attitude were made accountable to the users of the program, they tended to search more information, especially access a higher percentage of different cells than when they were not accountable. The additionally searched items were not primarily user

items, however, as the percentage of searched user information did not differ significantly between the 'no accountability' and 'accountability to users' conditions.

Variability of search was generally relatively high, indicating that participants tended to employ a non-compensatory decision strategy, also supported by the results for the Index of Compensatory Processing. Unlike expected, there were no significant differences between accountable and non-accountable participants in this respect. This suggests that task factors, for example, the complexity of the task, which was still relatively high with a choice set consisting of four alternatives differing on four attributes, overrode the effects of accountability. The only significant effect obtained was one of attitude on the variability of search across attributes for different cells. Participants with a pro-management attitude showed a significantly lower variability of search across attributes for different cells, indicating that they searched alternatives more equally and did not eliminate them from consideration as early as did participants with a pro-user attitude. This is consistent with the idea that participants with a pro-management attitude paid attention to both user and cost attributes; since alternatives had been constructed such that they were balanced in terms of their overall value, someone who was concerned with this overall value would find it hard to eliminate any of them from consideration.

The same attitude effect was observed when analysing the extent of compensatory processing as indicated by Koele and Westenberg's Index of Compensatory Processing and the Payne Index. Again, participants with a pro-management attitude showed a more compensatory and less attribute-wise search than participants with a pro-user attitude, lending further support to the conclusions above. Unlike expected, these measures were not affected by the type of accountability. Again, the effect of task complexity seems to have overridden that of the accountability manipulations.

Finally, the results obtained for the concentration of search measure did not show the predicted lower concentration of search on the chosen alternative for accountable compared to non-accountable participants, but revealed a (non-significant) tendency for the opposite effect for the total number of items searched, but not for the different cells searched. This suggests that when information was accessed repeatedly, accountable participants were more biased

towards their chosen alternative than non-accountable participants. However, this finding may also be interpreted as a stronger need to make sure that their preferred alternative was superior and as such would be consistent with the hypothesised stronger need for accuracy in these participants. A significant interaction between type of accountability and attitude furthermore suggested that only participants with a pro-user attitude were significantly affected by the type of audience they were accountable to. When they had been made accountable to the management, participants with a pro-user attitude showed a higher concentration of search on their chosen alternative than participants with a pro-management attitude. Participants with a pro-management attitude, on the other hand, did not show the same effect when they had been made accountable to the users of the program. This suggests that the conflict between having a pro-user attitude and being accountable to the management was resolved differently than the conflict between having a pro-management attitude and being accountable to the users of the program. In the former case, participants seemed to have become defensive and not given in to the accountability pressure, whereas in the latter case participants seemed to have been happy to take into account the preferences of the audience they were accountable to. The results for information evaluation and integration are expected to shed more light on the validity of this assumption.

Information evaluation and integration

1. Choices

Participants' choices were classified in terms of whether they were user choices or management choices, depending on whether the chosen alternative fared well on user attributes or on cost attributes, and would therefore either be easy to justify to users or to the management of a company. Given that there was a clear preference for user attributes in the sample and the fact that the pure management alternative (i.e., the alternative with positive outcomes on the two cost attributes and negative outcomes on the two user attributes) was only chosen by one participant, it was decided to compare participants in terms of whether they had chosen the pure user alternative (positive outcomes on both user attributes and negative outcomes on both cost attributes) or any of the other alternatives. The latter will be

referred to as management alternatives and, apart from the pure management alternative, included the two compromise alternatives with one positive and one negative outcome on one user and one cost attribute, respectively. They maximised management values in terms of a cost-benefit trade-off. Table 7.5 presents the frequencies with which the user alternative and one of the management alternatives were chosen, as a function of participants' attitude and type of accountability.

Table 7.5. Frequency of user and management choices at the first decision, as a function of type of accountability and attitude.

	No Accountability		Accountability to Users		Accountability to Management		
Attitude	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Σ
User Alt.	17	8	8	4	3	2	42
Man. Alt.	7	16	1	3	4	7	38
Σ	24	24	9	7	7	9	80

The user alternative was chosen by 52% of the participants (42/80), a management alternative was chosen by 48% of them (38/80). As expected, participants who had not been made accountable to any group chose according to their own attitude (χ^2 (1) = 6.76, p = .009), participants with a user attitude chose the user alternative more often than a management alternative (17/24 = 78%) and participants with a management attitude chose a management alternative more often than the user alternative (16/24 = 67%). Participants who had been made accountable to the users of the program, however, chose the user alternative in the majority of cases (12/16 =75%), irrespective of their own attitude (χ^2 (1) = 2.12, *ns*), whereas participants who had been made accountable to the management chose one of the management alternatives in the majority of cases (11/16= 69%), again, irrespective of their own attitude (χ^2 (1) = 0.78, *ns*). Hence, a Chi-squared analysis testing the relationship between type of chosen alternative and type of accountability yielded a significant effect (χ^2 (2) = 6.15, p = .046).

2. Attribute weights

It had been predicted that the weight judgements of participants who had not been made accountable to any group would reflect their attitude, that is, participants with a pro-user attitude would place particular importance on user attributes and participants with a pro-management attitude would place particular importance on cost attributes, whereas participants who had been made accountable to either the users of the program or the management of the company would adjust their weight judgements in such a way that they supported the position of the group they had been made accountable to (high user attribute weights when accountable to users and high cost attribute weights when accountable to the management), irrespective of participants' own attitude. In order to test this hypothesis, participants' weight judgements for the four attributes contained in the first and in the second information set were normalised so that they added up to 1, and the sum of user attribute weights and cost attribute weights at the first elicitation (before the first decision) and the second elicitation (after the first decision) as well as the difference between the second and first elicitation was calculated for each participant. Only the results for the user attribute weights are shown in Table 7.6, the results for cost attribute weights are the complimentary values to 1.

The results are consistent with expectations in that participants who had been made accountable to the management assigned lower weights to user attributes from the *first* information set than they had done initially, whereas participants who had not been made accountable hardly changed their weight judgements. Unlike predicted, however, participants who had been made accountable to the users of the program did not increase their user attribute weights but decreased them, particularly participants with a pro-user attitude. The most likely explanation for this is a regression effect, as these participants had assigned a very high weight to user attributes initially. A two-factorial ANOVA, testing the observed change in user attribute weights from the first to the second elicitation for any effects of type of accountability and attitude, did not result in any significant effects.

Table 7.6. Mean sum of user attribute weights for information set 1 and 2 before and after the first decision, as a function of type of accountability and attitude.

	Accountability					
	No		To Users		To the Management	
	Pro User (n=24)	Pro Man. (n=24)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)
sum of user attribute weights, set1						
after first decision	.732	.642	.744	.743	.609	.622
before first decision	.749	.640	.793	.735	.678	.668
difference	-.018	.002	-.049	.008	-.068	-0.45
sum of user attribute weights, set2						
after first decision	.742	.629	.753	.694	.738	.655
before first decision	.747	.598	.730	.642	.726	.642
difference	-.005	.031	.024	.053	.012	.013

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.

Interestingly, the results for the user attributes contained in the *second* information set, for which participants had also provided weight judgements twice at this point, but for which attribute outcomes had not been linked to certain alternatives yet, showed a different pattern. Participants who had been made accountable to the management did not decrease their user attribute weight judgements but increased them slightly. The same was true for participants in the 'no accountability' and 'accountability to users' condition, except for participants with a pro-user attitude who had not been made accountable, who showed no change. Accordingly, the ANOVA testing the effects of type of accountability and attitude on the change from the first to the second elicitation did not show any significant effects. This suggests that the importance differentiation process is employed to enhance particular alternatives with particular outcomes on particular attributes rather than attributes in general.

The sums of normalised user attribute weights obtained after the first decision were also subjected to two-factorial ANOVAs, testing the effects of attitude and type of accountability. The only significant effect in the analysis performed on weights for user

attributes from the *first* information set was a main effect of type of accountability ($F(2, 74) = 6.37$, $p = .003$). Post-hoc comparisons showed that, as expected, when participants were accountable to users, they assigned significantly higher weights to user attributes than when they were accountable to the management ($M = 0.74$ vs. $M = 0.62$, $t(79) = 3.56$, $p = .001$). Also, participants who were not accountable gave significantly higher user attribute weights than participants who were accountable to the management ($M = 0.69$ vs. $M = 0.62$, $t(79) = 2.43$, $p = .017$). Although the interaction effect between type of accountability and attitude did not reach conventional significance levels ($F(2,74) = 2.13$, $p = .126$), simple main effects analyses showed that the effect of attitude depended on accountability condition ($F(3,74) = 3.15$, $p = .030$). In line with expectations, a significant difference between participants with a pro-user attitude and participants with a pro-management attitude was only obtained in the 'no accountability' condition. Participants with a pro-user attitude assigned higher weights to user attribute than participants with a pro-management attitude ($M = 0.73$ vs. $M = 0.64$, $t(79) = 3.07$, $p = .003$). Also, participants with a pro-user attitude gave significantly higher user attribute weights when they were not accountable or accountable to users, compared to when they were accountable to the management ($F(2,74) = 4.52$, $p = .014$, no vs. management: $M = 0.73$ vs. $M = 0.61$, $t(79) = 2.82$, $p = .006$; users vs. management: $M = 0.74$ vs. $M = 0.61$, $t(79) = 2.65$, $p = .010$). Participants with a pro-management attitude, accordingly, assigned significantly lower weights to user attributes when they were not accountable or accountable to the management than when they were accountable to the users of the program ($F(2,74) = 3.34$, $p = .041$, no vs. users: $M = 0.64$ vs. $M = 0.74$, $t(79) = 2.33$, $p = .023$; management vs. users: $M = 0.62$ vs. $M = 0.74$, $t(79) = 2.38$, $p = .020$). The results are presented in Figure 7.4.

When conducting the same analysis on the sum of user attribute weights for the user attributes contained in the *second* information set, as expected, the main effect of type of accountability was not significant. The attitude main effect was significant, however ($F(1,74) = 10.18$, $p = .002$), with participants giving higher user attribute weights when they had a pro-user attitude than when they had a pro-management attitude ($M = 0.74$ vs. $M = 0.65$). This lends further support to the assumption that an importance differentiation process is employed to enhance the perceived value of certain alternatives rather than attributes in general.

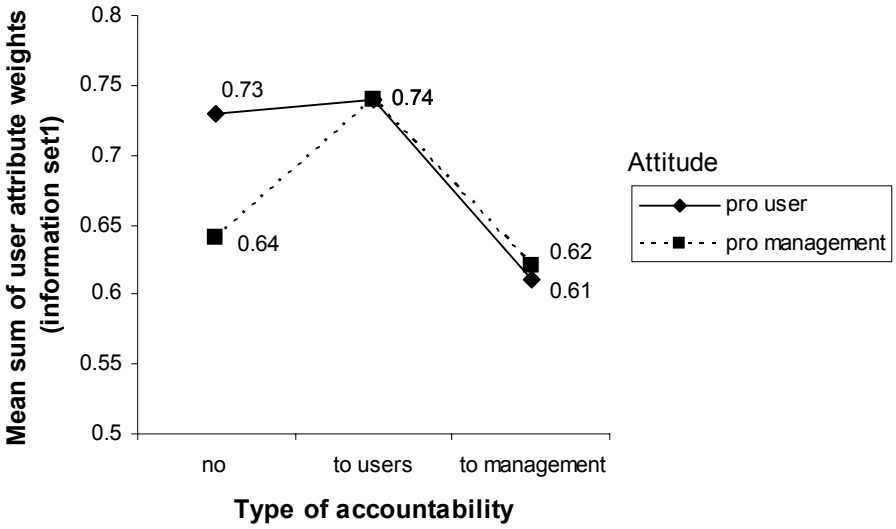


Figure 7.4. Mean sum of user attribute weights (information set1) after the first decision, as a function of type of accountability and attitude.

3. Overall evaluation of alternatives

Four types of dependent measures were analysed, (1) the difference between the suitability judgement for the chosen alternative and the average non-chosen alternative, (2) the difference between the MAUT prediction for the chosen alternative and the average non-chosen alternative, based on participants' unidimensional value and weight judgements after the first decision, (3) the difference between the suitability judgement for the user alternative and the average management alternative, and (4) the difference between the MAUT prediction for the user alternative and the average management alternative, again based on participants' unidimensional value and weight judgements after the first decision. As in the previous experiment, the MAUT model applied to calculate the overall evaluation of alternatives based on the attributes of the first information search only took the unidimensional values and weights of searched information into account (see Equation 6.5). The results are presented in Table 7.7.

Table 7.7. Measures of the evaluative difference between alternatives before and after the first decision, as a function of type of accountability and attitude.

	Accountability					
	No		To Users		To the Management	
	Pro User (n=24)	Pro Man. (n=24)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)
1. Rating chosen - average non-chosen alternative	44.86	32.22	35.04	41.19	35.24	30.44
2. MAUT prediction chosen - average non-chosen alternative						
set 1 attributes						
after first decision	30.16	21.12	34.64	26.74	35.86	12.80
before first decision	24.99	10.46	35.15	13.03	20.63	10.55
difference	5.18	10.67	-0.51	13.71	15.23	2.24
set 2 attributes						
after first decision	27.13	10.57	39.27	14.39	18.00	6.87
before first decision	28.99	13.65	39.28	10.82	16.10	-2.45
difference	-1.86	-3.07	-0.01	3.56	1.90	9.32
3. Rating user - average management alternative	32.08	15.28	29.11	25.95	18.10	2.30
4. MAUT prediction user - average man. alternative						
set 1 attributes						
after first decision	33.22	13.49	35.67	21.05	36.00	9.51
before first decision	28.75	6.44	31.80	12.45	23.83	8.49
difference	4.47	7.05	3.87	8.59	12.17	1.02
set 2 attributes						
after first decision	36.09	13.63	40.72	24.65	40.93	21.96
before first decision	37.75	10.27	39.26	15.31	42.69	19.10
difference	-1.66	3.36	1.46	9.34	-1.77	2.87

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.

3a. Rating chosen - average non-chosen alternative

It had been expected that participants who had been made accountable before the first decision would show a smaller difference between the suitability judgement for their chosen alternative and their average non-chosen alternative than participants who had not been made accountable, because this was an overt judgement. The judgmental difference between the chosen alternative and the average non-chosen alternative was also expected to be smaller for participants who had not adopted the preference of their external audience. This was assumed to be more likely for participants who had experienced a conflict due to the fact that their own attitude did not coincide with that of the group they had been made accountable to, resulting in an interaction between type of accountability and attitude.

In order to compare the 'no accountability' conditions against the conditions in which participants had been made accountable, a 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVA was conducted. Contrary to expectations, there was no significant main effect of accountability. There was, however, a significant main effect of attitude ($F(1,76) = 4.96$, $p = .029$) and a significant interaction between accountability and attitude ($F(1,76) = 4.99$, $p = .028$). The difference between the chosen alternative and the average non-chosen alternative was larger for participants with a pro-user attitude than for participants with a pro-management attitude ($M = 40.97$ vs. $M = 33.39$). However, this attitude main effect was qualified by the significant accountability-by-attitude interaction. The difference between accountability conditions was only significant for participants with a pro-user attitude ($F(1,76) = 5.91$, $p = .017$), and the effect of attitude was only significant in the 'no accountability' condition ($F(2,76) = 6.22$, $p = .003$). Within the group of participants with a pro-user attitude, those who were accountable, as expected, showed a smaller evaluative difference between the chosen alternative and the average non-chosen alternative than those who were not accountable ($M = 35.13$ vs. $M = 44.86$, $t(79) = 2.43$, $p = .017$), whereas there was no significant difference for participants with a pro-management attitude ($M = 35.15$ vs. $M = 32.22$). Also, non-accountable participants with a pro-user attitude showed a significantly larger evaluative difference than participants with a pro-management attitude ($M = 44.86$ vs. $M = 32.22$, $t(79) = 3.53$, $p = .001$), but there was no significant difference between the two groups when they were accountable ($M = 35.13$ vs. $M = 35.15$). This suggested that for the

overt measure of suitability judgements, accountability had the expected effect of diminishing rather than enhancing the judgmental difference between the chosen and the average non-chosen alternative, but only for participants with a pro-user attitude.

The hypothesis predicting an interaction between type of accountability and attitude was tested with an ANOVA including these two factors. The predicted significant interaction between type of accountability and attitude was found ($F(2,74) = 3.55$, $p = .034$). A simple main effects analysis indicated that the effect of attitude depended on the type of accountability ($F(3,74) = 4.73$, $p = .005$). Contrary to expectations, however, the only significant difference found was for participants in the 'no accountability' condition ($t(79) = 3.53$, $p = .001$); non-accountable participants with a pro-user attitude showed a higher judgmental difference between their chosen and average non-chosen alternative than non-accountable participants with a pro-management attitude ($M = 44.86$ vs. $M = 32.22$, $t(79) = 3.55$, $p = .001$). There was also a marginally significant effect of type of accountability for participants with a pro-user attitude ($F(2,74) = 2.99$, $p = .056$). As expected, the largest judgmental difference was observed in the 'no accountability' condition. It was significantly larger than the difference obtained in the 'accountability to users' condition ($M = 44.86$ vs. $M = 35.04$, $t(79) = 2.04$, $p = .045$), and marginally significantly larger than the difference obtained in the 'accountability to the management' condition ($M = 44.86$ vs. $M = 35.24$, $t(79) = 1.82$, $p = .072$). The observed differences between the suitability judgement for the chosen and the average non-chosen alternative are shown in Figure 7.5.

3b. MAUT prediction chosen - average non-chosen alternative

The differences between MAUT predictions for the chosen and the average non-chosen alternative had been predicted to show differentiation and consolidation effects, given that they were covert measures. Accordingly, it was expected that for accountable participants the evaluative difference between their chosen alternative and average non-chosen alternative would be larger than for participants who had not been made accountable. The MAUT predicted evaluative difference was also expected to be larger for participants who experienced a conflict due to the fact that their own attitude did not coincide with that of the

group they had been made accountable to, resulting in an interaction between type of accountability and attitude.

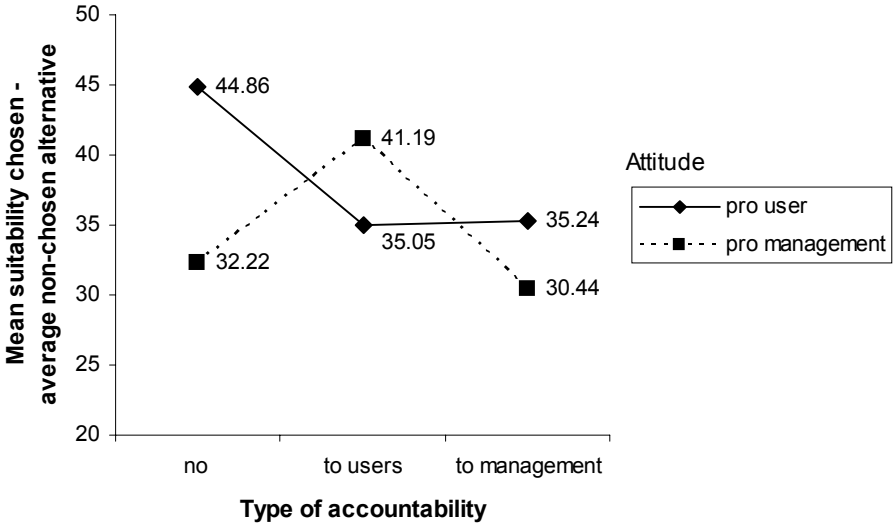


Figure 7.5. Mean difference in overall suitability between the chosen and average non-chosen alternative after the first decision, as a function of type of accountability and attitude.

An ANOVA was carried out to test the effect of accountability (yes vs. no) and attitude (pro-user vs. pro-management) on the *change* in the difference between the MAUT prediction for the chosen and the average non-chosen alternative from the first to the second elicitation, based on the four attributes contained in the *first* information set. Contrary to expectations, it did not yield any significant effects. The same analysis for the change in the difference between MAUT predictions based on the four parallel attributes contained in the *second* information set, however, did produce a significant main effect of accountability ($F(1,76) = 4.28$, $p = .042$). As expected, participants who had been made accountable showed a larger and positive evaluative change than participants who had not been made accountable ($M = 3.81$ vs. $M = -2.47$).

Similar analyses were performed on the differences between MAUT predictions for the chosen and the average non-chosen alternative obtained *after the first decision*. The ANOVA for the predictions based on the attributes in the *first* information set yielded a highly

significant main effect of attitude ($F(1,76) = 13.84$, $p = .006$). Participants with a pro-user attitude showed a significantly larger difference between the MAUT prediction for their chosen and average non-chosen alternative than participants with a pro-management attitude ($M = 32.17$ vs. $M = 20.23$). There was no significant interaction between accountability and attitude. The same was true for the analysis of the MAUT predictions based on the attributes in the *second* information set. The only significant effect obtained was a main effect of attitude ($F(1,76) = 14.10$, $p = .000$). Again, participants with a pro-user attitude showed a larger evaluative difference than participants with a pro-management attitude ($M = 28.27$ vs. $M = 10.41$).

In order to test the predictions regarding the effects of a conflict between the attitude of the group participants had been made accountable to and their own attitude, two-factorial ANOVAs testing the effects of type of accountability and attitude on the *change* in the MAUT predicted difference between the chosen and the average non-chosen alternative from the first to the second elicitation were conducted. For the MAUT predictions based on the attributes from the *first* information set, a marginally significant interaction between type of accountability and attitude was indeed obtained ($F(2,74) = 2.80$, $p = .068$). Simple main effect analyses suggested that for participants who had been made accountable to the users of the program, those with a pro-user attitude, as expected, tended to show a smaller (and, in fact, slightly negative) increase in the evaluative difference between their chosen and average non-chosen alternative, compared to participants with a pro-management attitude who would experience a stronger conflict and showed a large positive increase ($M = -0.51$ vs. $M = 13.71$, $t(79) = 1.68$, $p = .096$). Also, in line with predictions, for participants who had been made accountable to the management, the opposite tended to be true (although not significantly so); participants with a pro-user attitude tended to show a larger evaluative difference between their chosen and average non-chosen alternative than participants with a pro-management attitude ($M = 15.23$ vs. $M = 2.24$, see Figure 7.6).

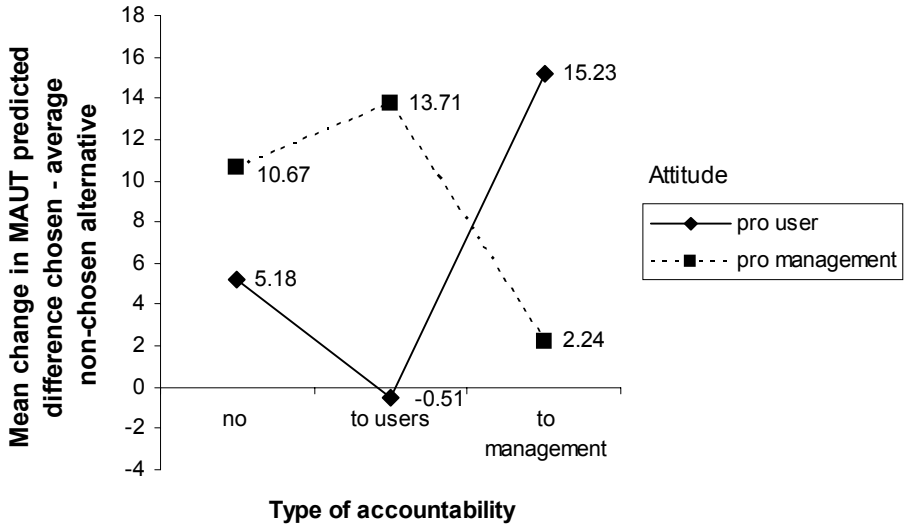


Figure 7.6. Mean change in the MAUT predicted evaluative difference between the chosen and average non-chosen alternative (based on the attributes from information set 1) from before to after the first decision, as a function of type of accountability and attitude.

The ANOVA for the MAUT predictions based on the attributes from the *second* information set yielded no significant effects. However, there was a weak tendency for a main effect of type of accountability ($F(2,74) = 2.33$, $p = .105$). In line with expectations, participants in the 'no accountability' condition showed the weakest differentiation (indeed a slight decrease in the difference between the MAUT prediction for their chosen and average non-chosen alternative, $M = -2.47$). Participants in the 'accountability to the management' condition showed the strongest differentiation ($M = 6.07$).

An ANOVA performed on the difference between the MAUT prediction for the chosen and the average non-chosen alternative *after the first decision*, based on the attributes from the *first* information set, only indicated a significant main effect of attitude ($F(1,74) = 12.42$, $p = .010$). As before, the predicted evaluative difference was larger for participants with a pro-user attitude than for participants with a pro-management attitude ($M = 32.17$ vs. $M = 20.23$). The same analysis for the MAUT prediction based on the attributes contained in the *second* information set also only showed a significant main effect of attitude ($F(1,76) =$

10.74, $p = .002$). For participants with a pro-user attribute, the difference between the MAUT prediction for the chosen and the average non-chosen alternative was larger than for participants with a pro-management attitude ($M = 28.27$ vs. $M = 10.41$). The main effect of type of accountability failed to reach standard significance levels ($F(2,74) = 1.89$, $p = .159$), but post-hoc contrasts suggested that the difference in MAUT predictions between the chosen and the average non-chosen alternative tended to be larger when participants had been made accountable to the users of the program than when they had been made accountable to the management of the company ($M = 28.39$ vs. $M = 19.34$, $p = .057$).

3c. Rating user - average management alternative

A large difference between the rating for the user and the average management alternative may be expected if individuals have clear user preferences and are not concerned with any benefits of alternatives regarding cost attributes. It had therefore been predicted that the judgmental difference between the user alternative and the average management alternative would be larger for participants with a pro-user attitude than with a pro-management attitude, and larger for participants who had been made accountable to the users of the program than for participants who had been made accountable to the management of the company, assuming that, on the whole, participants would adjust to the preferences of their audience. The difference had also been expected to be particularly small for participants with a pro-management attitude who had been made accountable to the management, and particularly large for participants with a pro-user attitude who had been made accountable to the users of the program, resulting in an interaction between type of accountability and attitude. This was because, in these cases, there was no conflict between participants' own preferences and those of the group they had been made accountable and both own preferences and choice would pull the ratings in the same direction. Participants who experienced a value conflict, on the other hand, were expected to succumb to the pressure created by the conflicting preferences of the group they had been made accountable to, and, for this overt measure of differentiation, demonstrate some adjustment to the preferences of their audience.

In order to test these hypotheses, a 3 (type of accountability: no vs. to users vs. to the management) by 2 (attitude: pro-user vs. pro-management) ANOVA was conducted. It

revealed a marginally significant main effect of attitude ($F(1,74) = 3.40$, $p = .069$). As expected, for participants with a pro-user attitude, the observed judgmental difference between the user and the management alternative tended to be larger than for participants with a pro-management attitude ($M = 28.97$ vs. $M = 14.23$). There was also a weak tendency for a main effect of type of accountability ($F(2,74) = 2.21$, $p = .117$). Again, as expected, for participants who had been made accountable to the management, the difference tended to be smaller than both for participants who had been made accountable to the users of the program ($M = 10.20$ vs. $M = 27.53$, $t(79) = 1.93$, $p = .058$) and for non-accountable participants ($M = 10.20$ vs. $M = 23.68$, $t(79) = 1.84$, $p = .070$). The expected interaction between type of accountability and attitude was not significant. However, as expected, participants who had been made accountable to the management showed a smaller judgemental difference between the user and average management alternative when they had a pro-management attitude than when they had a pro-user attitude ($M = 18.10$ vs. $M = 2.30$), whereas for participants who had been made accountable to the users of the program there was hardly any difference between those with a pro-user and those with a pro-management attitude ($M = 29.11$ vs. $M = 25.95$). The latter suggests that participants with a pro-management attitude who had been made accountable to the users of the program adjusted to the preferences of their audience easily. The mean differences between suitability ratings for the user and average management alternative after the first decision, which were observed in the different conditions, are shown in Figure 7.7.

3d. MAUT prediction user - average management alternative

Like for the actual suitability judgements, it had been expected that the evaluative difference between MAUT predictions for the user and the average management alternative would be larger for participants who had been made accountable to the users of the program than for participants who had been made accountable to the management of the company, with the 'no accountability' condition in-between, and larger for participants with a pro-user attitude compared to a pro-management attitude. In departure from the predictions for the suitability judgements, however, the difference had been expected to be particularly large for participants with a pro-user attitude who had been made accountable to the management and particularly small for participants with a pro-management attitude who had been made

accountable to the users of the program, indicating bolstering of a conflicting choice rather than adjustment to audience preferences.

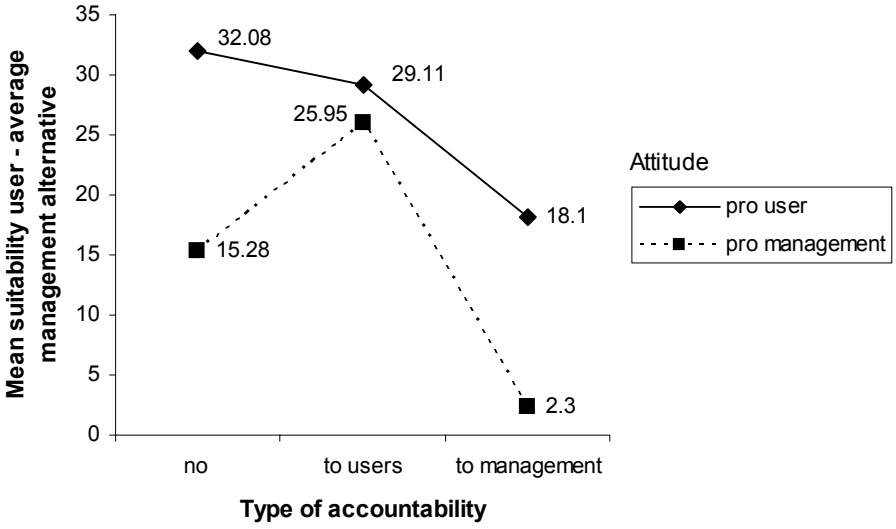


Figure 7.7. Mean difference in overall suitability between the user and average management alternative after the first decision, as a function of type of accountability and attitude.

These predictions were tested by ANOVAs including the 'type of accountability' factor and the 'attitude' factor. For the *change* in the difference between the MAUT prediction for the user and the average management alternative from before the first to after the first decision, based on the attributes from the *first* information set, no significant effects were obtained. The observed pattern of results was consistent with an interaction between type of accountability and attitude, although not in the expected direction; whereas for participants who had been made accountable to the users of the program, those with a pro-management attitude showed a *larger* (and not smaller) increase in the difference between the MAUT prediction for the user and average management alternative than those with a pro-user attitude (M = 8.59 vs. M = 3.87), the opposite was true for participants who had been made accountable to the management (M = 1.02 vs. M = 12.17). A large individual variability in responses prevented these differences from becoming significant, however.

The analysis of change in the difference between the MAUT prediction for the user and the average management alternative, calculated for the parallel attributes from the *second* information set, only revealed a marginally significant main effect of attitude ($F(1,74) = 3.31$, $p = .073$). Unlike expected, participants with a pro-management attitude tended to show an increase in the evaluative difference between the user and average management alternative from before the first to after the first decision, whereas participants with a pro-user attitude showed a slight decrease ($M = 4.30$ vs. $M = -0.98$).

The ANOVA that was performed on the MAUT predicted difference between the user and average management alternative *after the first decision*, based on the attributes contained in the *first* information set, only yielded the expected main effect of attitude, which was highly significant ($F(1,74) = 20.29$, $p = .000$). As predicted, participants with a pro-user attitude showed a larger evaluative difference between the user and average management alternative than participants with a pro-management attitude ($M = 34.26$ vs. $M = 13.91$). The expected interaction between type of accountability and attitude was not significant. Although participants with a pro-user attitude who had been made accountable to the management, as expected, showed a much larger difference between the MAUT prediction for the user and average management alternative ($M = 36.00$ vs. $M = 9.51$), a large individual variability in judgements again meant that this difference was not significant.

The analysis for the MAUT predictions based on the attributes from the *second* information set revealed a marginally significant main effect of type of accountability ($F(2,74) = 2.41$, $p = .097$). Post-hoc comparisons indicated a marginally significant difference between the 'no accountability' and the 'accountability to users' condition. For participants who had been made accountable to the users of the program the predicted difference tended to be larger than for participants who had not been made accountable ($M = 33.69$ vs. $M = 24.86$, $t(79) = 1.87$, $p = .065$). Unexpectedly, though, the mean predicted difference obtained in the 'accountability to users' condition was not significantly larger than the mean predicted difference obtained in the 'accountability to the management' condition ($M = 33.69$ vs. $M = 30.26$). In addition, there was again a highly significant main effect of attitude ($F(1,74) = 26.91$, $p = .000$). For participants with a pro-user attitude, the predicted difference was larger

than for participants with a pro-management attitude (M = 37.98 vs. M = 17.43). The interaction between type of accountability and attitude was not significant.

Summary of the results for information evaluation and integration at the first decision

The results relating to information evaluation and integration at the first decision, although they did not always exactly match the predictions, were largely consistent with them. Overt measures, such as participants' choices, weight judgements and suitability ratings for the user compared to the average management alternative, indicated some adjustment to the inferred values of the audience participants had been made accountable to. As expected, participants who had been made accountable to the users of the program were more likely to choose the user alternative and participants who had been made accountable to the management were more likely to choose one of the management alternatives, no matter what their own attitude was, while the choices of participants in the 'no accountability' condition depended on their attitude; they were more likely to choose the user alternative if their attitude was user-oriented and more likely to choose one of the management alternatives if their attitude was more management-oriented.

Likewise, the weight judgements for user attributes by participants who had been made accountable to the users of the program were significantly higher than those by participants who had been made accountable to the management, irrespective of participants' own attitude. The weight judgements by participants who had not been made accountable, on the other hand, corresponded with their attitude. Interestingly, this effect was only observed for the attributes contained in the first information set and not for those in the second, for which attribute outcomes had not yet been linked to certain alternatives at the time the judgements were made. Weight judgements for the attributes contained in the second information set only showed an effect of attitude; participants with a pro-user attitude assigned larger weights to user attributes than participants with a pro-management attitude. Whereas the adjustment effect was highly significant in participants' weight judgements after the first decision, the change in weight judgements, calculated as the difference between the weights obtained at the second elicitation (after the first decision) and the weights obtained at the first elicitation

(before any experimental manipulations), although showing the predicted trend, was not. It may therefore be better to speak of importance *adjustment* rather than *differentiation* in this case. These findings also suggest that the problem of which dependent variable one chooses to inspect is not a trivial one.

The overall differentiation between user and management alternatives, as captured in participants' suitability ratings for the user compared to the average management alternative, only showed a weak tendency for adjustment to the audience participants had been made accountable to, but, as expected, the observed judgmental difference between the user and average management alternative was smaller for participants who had been made accountable to the management than for participants who had either been made accountable to the users of the program or participants who had not been made accountable. If there was a conflict between participants' own attitude and the preferences of the group they had been made accountable to, adjustment seemed to have been stronger for participants with a pro-management attitude who had been made accountable to the users of the program than the opposite combination.

The difference between MAUT predictions for the user compared to the average management alternative, as a covert measure of differentiation, had been expected to show bolstering of participants' own preferences rather than adjustment to audience preferences when they experienced a conflict between their own preferences and those of the audience. However, this tendency was not consistently evident, neither in the change of the evaluative difference from before the first to after the first decision, nor in the evaluative difference obtained after the first decision. Rather, the observed change in the difference between the MAUT prediction for the user and the average management alternative from the first to the second elicitation (based on the attributes from the first information set), although not significantly so, again suggested that participants with a pro-management attitude who had been made accountable to the users of the program adjusted to the preferences of their audience, whereas participants with a pro-user attitude who had been made accountable to the management bolstered their own preference, as both showed a relatively large increase in the evaluative difference between the user and average management alternative. The same measure, but based on the parallel attributes from information set 2, showed a marginally

significant main effect of attitude such that participants with a pro-management (and not a pro-user!) attitude showed the larger increase in the evaluative difference between the user and average management alternative, mainly because participants with a pro-management attitude who had been made accountable to the users of the program showed a particularly large increase. This seems to support the assumption that they easily gave in to user pressures.

The analyses performed on the differences between MAUT predictions for the user and the average management alternative after the first decision mainly yielded the expected effects of participants' own attitude; participants with a pro-user attitude showed a larger evaluative difference between the user and average management alternative than participants with a pro-management attitude, both for the attributes of the first and of the second information set. The analysis carried out on the parallel attributes also indicated the expected, marginally significant tendency for participants who had been made accountable to the users of the program to show a larger evaluative difference between the user and average management alternative than participants who had not been made accountable. In order for such an effect to be obtained, participants with a pro-management attitude who had been made accountable to the users of the program must have shown an adjustment to the preferences of their audience.

The most notable tendency emerging from the analyses testing the actual and predicted evaluative differences between the chosen and average non-chosen alternative was a main effect of attitude. Especially the analyses investigating the predicted evaluative differences between the chosen and average non-chosen alternative after the first decision showed this effect to be the strongest and often the only significant one. Generally, the obtained evaluative differences were larger for participants with a pro-user attitude than for participants with a pro-management attitude. This finding may be explained by another effect that has already been discussed earlier, namely that participants with a pro-user attitude seemed to have only paid attention to user attributes, whereas participants with a pro-management attitude were concerned with both user and cost attributes. This meant that participants with a pro-user attitude who chose the user alternative rated the overall value of this alternative as distinctly higher than the overall value of the management alternatives, resulting in a large evaluative difference between the chosen and average non-chosen alternative, whereas for participants

with a pro-management attitude the difference was smaller, because they considered an alternative with advantages on both user attributes as still relatively good. Indeed, there was no significant difference in the suitability rating of the user alternative by participants with a pro-user attitude and a pro-management attitude (M = 80.25 vs. M = 74.13, $t(78) = 1.36$, p = .176), whereas there was a significant difference in their rating of the pure management alternative (M = 37.50 vs. M = 48.00, $t(78) = 2.35$, p = .021). The high median obtained for the attitude measure, indicating that even participants with a pro-management attitude highly valued the user alternative, also supports this assumption.

This effect has important implications for the interpretation of the evaluative difference between the chosen and average non-chosen alternative observed in the various conditions. A large difference after the first decision does not necessarily only signal a strong differentiation between the chosen and the non-chosen alternatives, but also that participants clearly preferred the user alternative to the management alternatives. A small difference, on the other hand, may mean that participants valued user and cost aspects and therefore the user and the management compromise alternatives equally. The same applies to those dependent measures that analysed the *change* in the evaluative difference between the chosen and average non-chosen alternative from before to after the first decision; an increase may signal that preferences have shifted from a management to the user alternative, whereas a decrease may signal that preferences have shifted from the user to one of the management alternatives. Ideally, one would have included the type of chosen alternative as another factor in the analysis, but this was not possible, due to the fact that the sample size was too small and that when participants did not experience any value conflict, they rarely chose an alternative not in accordance with their own attitude.

The analyses performed on the difference between the suitability rating for the chosen and average non-chosen alternative, apart from an attitude main effect, yielded an interaction between accountability and attitude, such that accountable participants showed a smaller evaluative difference between the chosen alternative and the average non-chosen alternative than non-accountable participants, but only if they had a pro-user attitude. For participants with a pro-management attitude, no significant difference was observed. Accountability had been expected to reduce the evaluative difference between the chosen and average non-chosen

alternative, given that this was an overt measure of differentiation. The reason why this effect was not obtained for participants with a pro-management attitude was that, when they had been made accountable to the users of the program, they showed an enlarged evaluative difference between the chosen and average non-chosen alternative, thereby increasing the mean for the accountability condition and making it not significantly different from the mean observed for the 'no accountability' condition. Participants with a pro-user attitude, on the other hand, did show a significantly smaller evaluative difference between the chosen and average non-chosen alternative when they had been made accountable, even when they had been made accountable to an audience with compatible preferences.

The interaction effect remained significant when the effects of the particular types of accountability rather than accountability per se were analysed. It was caused, however, not by different response patterns in the 'accountability to users' and accountability to the management' conditions, but by the fact that in the 'no accountability' condition, participants with a pro-user attitude showed a higher judgmental difference between their chosen and average non-chosen alternative than those with a pro-management attitude. When participants had been made accountable to either the users of the program or the management of the company, no significant differences between participants with different attitudes emerged. This suggests that, rather than resisting the preferences of the audience and defending their deviant choice, participants showed some adjustment to these preferences, as already evident in choices and weight judgements.

The analyses of the MAUT predicted difference between the chosen and the average non-chosen alternative, however, as expected, told a slightly different story. The change in the MAUT predicted difference between the chosen and the average non-chosen alternative, based on the attributes from the first information set, showed a marginally significant interaction between type of accountability and attitude. Among participants who had been made accountable to the users of the program, those with a pro-management attitude tended to increase the evaluative difference between their chosen and average non-chosen alternative more strongly than those with a pro-user attitude, whereas the opposite tended to be true for participants who had been made accountable to the management. This suggests bolstering of the chosen alternative, whenever a value conflict existed. A closer inspection of the means

observed in these conditions, depending on whether participants solved the value conflict in favour of the audience preferences and chose an alternative consistent with them, or in favour of their own attitude, that is, chose an alternative inconsistent with audience preferences, supports this assumption. Participants with a pro-management attitude who had been made accountable to the users of the program showed a larger evaluative difference between their chosen and average non-chosen alternative if they had chosen according to their own attitude rather than according to audience preferences (M = 22.89, n = 3 vs. M = 6.83, n = 4). The same was true for participants with a pro-user attitude who had been made accountable to the management of the company (M = 17.77, n = 3 vs. M = 13.33, n = 4). Given the small ns, it is not surprising that in the ANOVA performed to analyse these differences, the expected interaction between type of accountability (to users vs. to management) and type of chosen alternative (user vs. management) was not significant (F (1,10) = 1.15, p = .308). The trends in the data seem to be rather clear, however.

Supporting this assumption is the fact that a similar pattern was observed for the differences between MAUT predictions for the chosen and average non-chosen alternative after the first decision, based on the attributes from the first information set. Although the interaction between type of accountability and attitude was not significant, simple main effect analyses suggested that when participants had been made accountable to the management of the company, the MAUT predicted evaluative difference between the chosen and the average non-chosen alternative was larger when participants had a pro-user attitude than when they had a pro-management attitude. There was no difference, however, when participants had been made accountable to the users. An inspection of the means in the value conflict conditions again suggested that the obtained difference was larger when participants had chosen an alternative that was inconsistent with the preferences of the group they had been made accountable to than when they had chosen an alternative that was consistent with these preferences (accountability to users: M = 39.97 vs. M = 16.81, accountability to the management: M = 43.05 vs. M = 30.47). This time the expected interaction between type of accountability and type of chosen alternative approached significance (F (1,10) = 4.11, p = .070).

Also, as expected, MAUT predictions based on the parallel attributes, which had not been linked to alternatives yet, did not show these effects, neither the change in the evaluative difference between the chosen and average non-chosen alternative from before the first to after the first decision, nor the difference obtained after the first decision. In both cases, analyses of the value conflict conditions only showed an effect of the type of chosen alternative (change: F (1,10) = 3.63, p = .086, after first decision: F (1,10) = 37.27, p = .000). For participants who had chosen the user alternative, a larger evaluative difference was obtained than for participants who had chosen a management alternative (change: M = 12.05 vs. M = -6.58, after the first decision: M = 43.55 vs. M = -11.16). The ANOVAs performed on the whole sample, accordingly, did not yield any significant effects, apart from the above mentioned attitude main effects.

Finally, the comparison of the MAUT predicted evaluative difference between the chosen and average non-chosen alternative for accountable and non-accountable participants did not yield any consistent results. The change in the difference between the MAUT prediction for the chosen and the average non-chosen alternative from before the first to after the first decision, based on the four attributes contained in the first information set, did not show the expected larger increase for accountable compared to non-accountable participants, but the change in the predictions based on the four parallel attributes contained in the second information set, which had not been expected to show the effect, did. Whereas participants who had been made accountable showed an overall increase in the evaluative difference between the chosen and average non-chosen alternative, participants who had not been made accountable showed a decrease.

The analyses performed on the differences between MAUT predictions for the chosen and the average non-chosen alternative obtained after the first decision only yielded significant main effects of attitude, both for predictions based on the attributes from the first and the second information set. Again, participants with a pro-user attitude showed a significantly larger difference between the MAUT predictions for their chosen and average non-chosen alternative than participants with a pro-management attitude.

Final decision

Information search

The same measures as for the first information search were analysed. Table 7.8 shows the results obtained in each condition. In addition, given the redundancy of the information items contained in the second information set, the final information search also allowed to test the extent to which information search was biased to support the chosen alternative, as captured by the Biased Search Index outlined in Chapter 6. The index takes into account both the expected value of searched information and its importance, by comparing the mean product of unidimensional values and their corresponding weights for the information items searched for a particular alternative to the mean product that would be expected if information search was completely balanced. For an unbiased information search, a value of 1 is expected. A value > 1 indicates that information search was biased to support the searched alternative and a value < 1 indicates that it was biased to devalue the searched alternative. An overall measure of confirmation bias for the chosen alternative compared to the non-chosen alternatives can be computed by dividing the BSI for the chosen alternative by the average BSI for the non-chosen alternatives. A value > 1 is expected if there is a bias to support the chosen over the average non-chosen alternative, a value < 1 if there is a bias to support the average non-chosen alternative over the chosen alternative.

1. Depth of search

Overall, the final information search was less exhaustive than the first information search, both in terms of the percentage of total number of items searched and the percentage of different cells searched (total number of items searched: 101% vs. 115%, $t(79) = 2.84$, $p = .006$, two-sided, different cells searched: 76% vs. 83%, $t(79) = 3.57$, $p = .001$, two-sided). This may have been expected, given that participants had already established a preference for one of the alternatives when they made their first decision, and given the fact that the information contained in the second search set was largely redundant with the information contained in the first set.

It had been predicted that accountable participants would search more information than participants who had not been made accountable. In order to compare the control group to the experimental groups, two-factorial ANOVAS testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro-management) were conducted. In line with expectations, there was a marginally significant main effect of accountability on the percentage of total number of items searched as well as different cells searched (total number of items searched: $F(1,76) = 3.76$, $p = .056$, different cells searched: $F(1,76) = 3.70$, $p = .058$). Accountable participants tended to search a higher percentage of information than non-accountable participants (total number of items searched: $M = 106\%$ vs. $M = 81\%$, different cells searched: $M = 78\%$ vs. $M = 67\%$). In both cases, there was also a significant main effect of attitude (total number of items searched: $F(1,76) = 4.19$, $p = .044$, different cells searched: $F(1,76) = 6.61$, $p = .012$). Like during the first information search, participants with a pro-management attitude searched a higher percentage of information than participants with a pro-user attitude (total number of items searched: $M = 110\%$ vs. $M = 92\%$, different cells searched: $M = 82\%$ vs. $M = 70\%$).

The control group was also compared to the experimental groups in ANOVAs which, apart from the attitude factor, either included the 'time of accountability' factor or the 'type of accountability' factor. It had been expected that participants who had been made accountable before the final information search would search a higher percentage of information than participants who had been made accountable before the first search, resulting in a significant main effect of time of accountability. The analyses only showed a weak tendency for such an effect (total number of items searched: $M = 108\%$ vs. $M = 103\%$, $F(2,74) = 2.07$, $p = .134$, different cells searched: $M = 80\%$ vs. $M = 76\%$, $F(2,74) = 2.15$, $p = .124$). The main effects of attitude were significant again, however (total number of items searched: ($F(1,74) = 4.18$, $p = .044$), different cells searched: ($F(1,74) = 7.81$, $p = .007$). Also, for the percentage of total number of items searched, the main effect of time of accountability was qualified by a marginally significant interaction between time of accountability and attitude ($F(2,74) = 2.97$, $p = .057$). Simple main effect analyses indicated that the main effect of time of accountability was only significant for participants with a pro-user attitude ($F(2,74) = 4.12$, $p = .020$), and that the effect of attitude depended on time of accountability ($F(3,74) = 3.05$, $p = .034$). Participants with a pro-user attitude searched a significantly higher percentage of information

when they had been made accountable before the final decision than when they had not been made accountable (M = 114% vs. M = 62%, $t(79) = 2.75$, $p = .008$). There was also a marginally significant difference between the 'accountability before the final decision' and the 'accountability before the first decision' conditions (M = 114% vs. M = 85%, $t(79) = 1.88$, $p = .064$). Furthermore, when participants had either not been made accountable or when they had been made accountable before the first decision, those with a pro-management attitude searched a higher percentage of information than those with a pro-user attitude (no accountability: M = 101% vs. M = 62%, $t(79) = 1.77$, $p = .080$, accountability before the first decision: M = 121% vs. M = 85%, $t(79) = 2.33$, $p = .022$).

Based on the results for the first information search, it was also expected that participants who had been made accountable to the management would search a higher percentage of information than participants who had either been made accountable to the users of the program or not been made accountable. The expected main effect of type of accountability was marginally significant for the percentage of total number of items searched ($F(2,74) = 2.57$, $p = .084$), and was significant for the percentage of different cells searched ($F(2,74) = 3.17$, $p = .048$). Post-hoc simple contrasts indicated that non-accountable participants searched a lower percentage of different cells than participants who had been made accountable to the users of the program (total number of items searched: 81% vs. 112%, $p = .028$, different cells searched: 67% vs. 82%, $p = .020$). No other comparison was significant. Again, the attitude main effects were significant (total number of items searched: $F(1,74) = 4.21$, $p = .044$, different cells searched: $F(1,74) = 8.19$, $p = .005$).

Three-factorial ANOVAs were conducted in order to test the joint effects of time of accountability (before the first decision vs. before the final decision), type of accountability (to users of the program vs. to the management of the company) and attitude (pro-user vs. pro-management) on the percentage of searched information. The control group was not included in these analyses, since it was not part of a full factorial design of these factors. The ANOVA performed on the percentage of total number of items searched did not yield the expected main effect of time of accountability, but did show a significant interaction between time of accountability and attitude ($F(1,56) = 4.16$, $p = .046$).

Table 7.8. Process measures for the final information search, as a function of time of accountability, type of accountability and attitude.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
1. Depth of search (in %)										
total # of items	62	101	96	117	71	124	120	117	108	88
different cells	58	77	74	84	57	88	83	88	75	74
2. Searched user information (in %)										
total # of items	64	55	65	63	73	62	61	58	65	59
different cells	63	51	60	61	69	55	52	52	62	54
3. Variability of search										
across alternatives										
total # of items	1.33	1.58	1.39	0.91	1.44	1.02	1.12	1.57	1.61	1.54
different cells	1.17	0.79	0.70	0.41	1.06	0.55	0.59	0.69	0.66	1.06
across attributes										
total # of items	0.96	1.14	1.59	1.35	1.22	1.78	1.94	1.08	1.25	0.99
different cells	0.80	0.66	0.80	0.66	0.80	0.41	0.31	0.19	0.59	0.53
4. Compensatory processing	0.28	0.52	0.52	0.68	0.30	0.67	0.64	0.62	0.56	0.37

Table 7.8. cont.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
5. Pattern of search (Payne Index)	-0.46	-0.12	-0.42	-0.56	-0.23	-0.35	-0.54	-0.32	-0.66	-0.38
6. Concentration of search										
total # of items	0.78	0.36	0.48	0.16	0.68	0.16	0.35	0.10	0.48	0.35
different cells	0.62	0.29	0.34	0.13	0.68	0.17	0.31	0.15	0.33	0.30
7. Confirmation bias (BSI)										
chosen alternative	1.14	1.04	1.17	1.19	1.14	1.14	1.37	0.87	1.50	1.09
average non-chosen alternative	0.89	1.00	1.10	1.07	1.13	1.00	1.01	1.09	0.99	1.08
chosen alt./ average non-chosen alt.	1.55	1.05	1.30	1.23	1.36	1.21	0.82	0.82	1.61	1.06

A simple main effects analysis showed that participants with a pro-management attitude searched a higher percentage of information items than participants with a pro-user attitude when they had been made accountable before the first decision (M = 121% vs. M = 83 %, $F(1,56) = 4.79$, $p = .033$), but not when they had been made accountable before the final decision (M = 102% vs. M = 114%, see Figure 7.8). The same interaction between time of accountability and attitude was obtained for the percentage of different cells searched, although it was only marginally significant in this case ($F(1,56) = 2.93$, $p = .092$). Nevertheless, there was a highly significant difference in the percentage of different cells searched between participants with different attitudes when they had been made accountable before the first decision only ($F(1,56) = 7.46$, $p = .008$). Participants with a pro-management attitude searched a higher percentage of different cells than participants with a pro-user attitude (M = 86% vs. M = 67%); this difference was not significant for participants who had only been made accountable before the final decision (M = 81% vs. M = 79%). In addition to the interaction between time of accountability and attitude, there was a significant main effect of attitude for the percentage of different cells searched ($F(1,56) = 4.69$, $p = .035$). Participants with a pro-management attitude searched a higher percentage of different cells than participants with a pro-user attitude (M = 84% vs. M = 73%).

2. Searched user information

The overall percentage of searched user information was 62% when related to the total number of items searched. This was higher than for the first information search (60%), the difference was marginally significant ($t(79) = 1.97$, $p = .052$, two-sided). The overall percentage of searched user information related to different cells searched was also higher than the value obtained for the first search, but not significantly so (M = 58% vs. M = 56%).

It had been expected that only participants who had been made accountable before the first decision would adjust the type of the information they searched to the group they had been made accountable to, that is, would search relatively more or less user information, depending on whether they had been made accountable to the users of the program or to the management of the company. Participants who had only been made accountable before the

final decision, on the other hand, would search information according to their own attitude, resulting in a significant interaction between time of accountability and type of accountability and time of accountability and attitude.

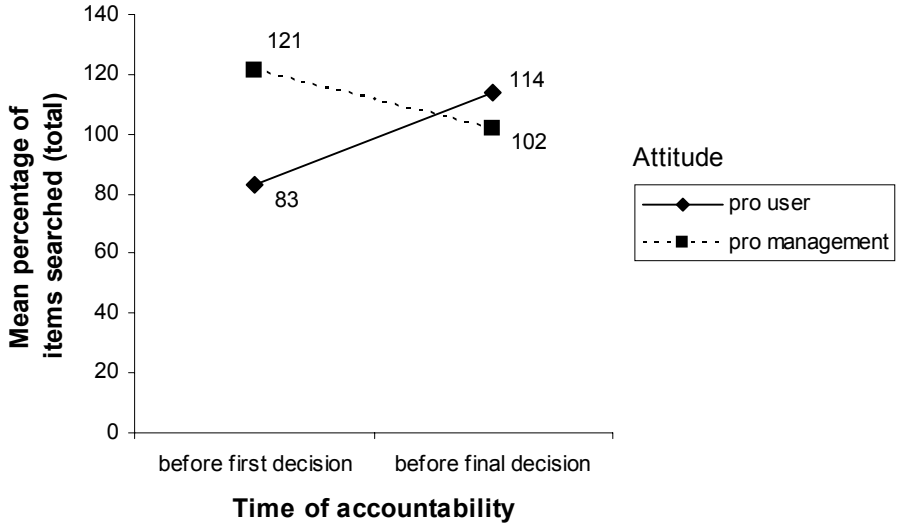


Figure 7.8. Mean percentage of items searched during the final information search (based on the total number of items searched), as a function of time of accountability and attitude.

Three-factorial ANOVAs performed to test these hypotheses revealed a marginally significant main effect of time of accountability on the percentage of searched user information calculated for the total number of items searched ($F(1,56) = 3.28$, $p = .076$), and a significant main effect of time of accountability on the percentage of searched user information calculated for different cells searched ($F(1,56) = 4.03$, $p = .049$). They also showed marginally significant main effects of attitude for both measures (total number of items searched: $F(1,56) = 3.89$, $p = .053$, different cells searched: $F(1,56) = 3.66$, $p = .061$). Participants who had been made accountable before the first decision tended to search a higher percentage of user information than participants who had been made accountable before the final decision (total number of items searched: $M = 66\%$ vs. $M = 61\%$, different cells searched: $M = 61\%$ vs. $M = 55\%$). Also, in line with expectations, participants with a pro-user attitude tended to search a higher percentage of user information than participants

with a pro-management attitude (total number of items searched: M = 66% vs. M = 60%, different cells searched: M = 61% vs. M = 55%). In addition, for the percentage of user information calculated for different cells searched, a marginally significant interaction between type of accountability and attitude was obtained ($F(1,56) = 3.92$, $p = .053$). A simple main effects analysis showed that the effect of type of accountability was only significant for participants with a pro-user attitude ($F(1,56) = 5.48$, $p = .023$), and that the effect of participants' attitude was only significant when they had been made accountable to the management of the company ($F(1,56) = 7.58$, $p = .008$). Participants with a pro-user attitude searched more user information when they had been made accountable to the management of the company than when they had been accountable to the users of the program (M = 66% vs. M = 57%), whereas there was no significant difference when participants had a pro-management attitude (M = 55% vs. M = 56%). Also, when participants had been made accountable to the management of the company, those with a pro-user attitude searched more user information than those with a pro-management attitude (M = 66% vs. M = 55%), whereas participants with a pro-management attitude who had been made accountable to the users of the program searched as much user information as participants with a pro-user attitude (M = 56% vs. M = 57%, see Figure 7.9). At first, these results seem to be at odds with the predictions, but may be explained if one assumes that participants with a pro-user attitude who thought they had to justify their decision to the management were under pressure, especially if they chose not to adjust to the opinion of the group they were made accountable to and preferred the user alternative instead. It may then be expected that they particularly concentrated on the positive attributes of this alternative, in order to be able to defend their choice. Participants who had been classified as having a pro-management attitude but were accountable to the users, on the other hand, seemed to have been much more willing to adjust their preference to that of the group they were accountable to. This is not surprising, given that attitudes were inclined towards user considerations anyway.

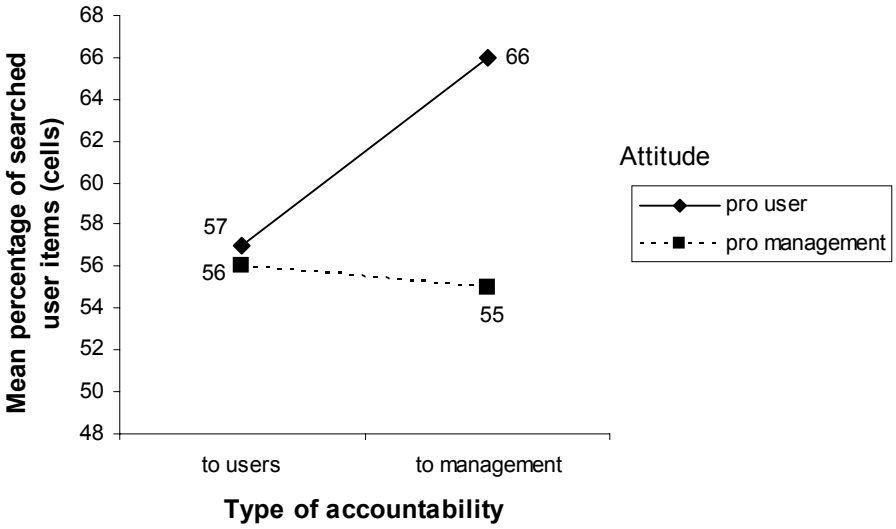


Figure 7.9. Mean percentage of user information searched during the final information search (based on different cells searched), as a function of type of accountability and attitude.

The percentage of searched user information calculated for the total number of items searched as well as different cells searched did not differ significantly between accountable and non-accountable participants, as the main effect of accountability in the 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVAs performed on these measures was not significant. For both measures, there was a significant main effect of attitude, however (total number of items searched: $F(1,76) = 5.48$, $p = .022$, different cells searched: $F(1,76) = 6.13$, $p = .015$). As expected, participants with a pro-user attitude searched a higher percentage of user information than participants with a pro-management attitude (total number of items searched: $M = 66\%$ vs. $M = 59\%$, different cells searched: $M = 61\%$ vs. $M = 55\%$).

Unexpectedly, the percentage of searched user information calculated for the total number of items searched as well as different cells searched did not vary as a function of type of accountability. The attitude main effect was significant again for both measures, however ($F(1,74) = 6.49$, $p = .013$, different cells searched: $F(1,74) = 7.26$, $p = .009$).

There was neither a significant main effect of time of accountability on the percentage of searched user information related to the total number of items searched nor on the percentage of searched user information related to different cells searched. Again, only the attitude main effects were significant (total number of items searched: $F(1,74) = 6.38$, $p = .014$, different cells searched: $F(1,74) = 6.85$, $p = .011$).

3. Variability of search

The overall mean variability of search across alternatives calculated for the total number of items searched was slightly lower than for the first information search ($M = 1.35$ vs. $M = 1.43$), but not significantly so. The opposite trend was observed for variability of search across alternatives calculated for different cells searched ($M = 0.76$ vs. $M = 0.65$, $t(79) = 1.76$, $p = .081$, two-sided). The latter supports the assumption that participants' information search concentrated on only a subset of the available alternatives. The overall mean variability of search across attributes was slightly larger for the final than for the first information search ($M = 1.34$ vs. $M = 1.30$ for total number of items searched and $M = 0.57$ vs. $M = 0.54$ for different cells searched), but not significantly so. The fact that the observed variability of search was larger when calculated for the total number of items searched than for different cells searched (across alternatives: $M = 1.35$ vs. $M = 0.76$, $t(79) = 7.22$, $p = .000$, across attributes: $M = 1.34$ vs. $M = 0.57$, $t(79) = 7.68$, $p = .000$) indicates that when participants searched information repeatedly, they tended to do this less evenly. Variability of search of different cells was significantly larger across alternatives than it was across attributes ($M = 0.76$ vs. $M = 0.57$, $t(79) = 2.31$, $p = .041$). The same difference was not significant where the total number of items was concerned, however ($M = 1.35$ vs. $M = 1.34$). The former effect again lends support to the assumption that participants did not search information for all alternatives; however, the alternatives they did consider, they tended to search a similar number of attributes for.

It had been expected that the variability of search across alternatives and attributes would be lower for accountable than for non-accountable participants. The observed variability of search across alternatives was indeed lower for accountable than for non-accountable participants (total number of items searched: $M = 1.33$ vs. $M = 1.45$, different

cells searched: M = 0.71 vs. M = 0.98), but in the ANOVAs performed to test this hypothesis the main effect of accountability was not significant. The variability of search across attributes showed the predicted trend when it was calculated for different cells searched (M = 0.53 vs. M = 0.73), but not when it was calculated for the total number of items searched (M = 1.41 vs. M = 1.05). In neither case were significant main effects of accountability obtained, however. There were also no significant attitude main effects or significant interactions between accountability and attitude observed in any of the analyses.

Furthermore, it had been expected that participants who had been made accountable before the decision would show a lower variability of search across alternatives and attributes than participants who had been made accountable before the first decision, resulting in significant main effects of time of accountability. Two-factorial ANOVAs testing the effects of time of accountability (no vs. before the first decision vs. before the final decision) and attitude (pro-user vs. pro-management) were carried out to test these predictions. There was indeed a marginally significant effect of time of accountability on the variability of search across attributes calculated for different cells searched (F(2,74) = 2.56, p = .084), but not for any of the other measures. Post-hoc simple contrasts suggested that, as predicted, participants who had been made accountable before the final decision tended to show a lower variability of search across attributes than both non-accountable participants (M = 0.41 vs. M = 0.73, p = .056) and participants who had been made accountable before the first decision (M = 0.41 vs. M = 0.66, p = .068). For variability of search across alternatives calculated for different cells searched, a weak tendency for an interaction between time of accountability and attitude was observed (F(2,74) = 2.34, p = .103). Simple main effects analyses suggested that the effect of time of accountability tended to be limited to participants with a pro-user attitude (F(2,74) = 2.03, p = .138). When they had been made accountable before the final decision, they showed a significantly lower variability of search across alternatives calculated for different cells searched than when they had not been made accountable (M = 0.62 vs. M = 1.17, t(79) = 2.00, p = .049). The difference to participants who had been made accountable before the first decision was in the predicted direction, but not significant (M = 0.62 vs. M = 0.86). The 'type of accountability' main effect did not reach standard significance levels in the ANOVAs performed on any of the variability of search measures.

Three-factorial ANOVAs performed on measures of variability of search across alternatives and attributes only yielded significant effects when variability of search was related to different cells searched but not when it was related to the total number of items searched. For variability of search across alternatives, there was a significant interaction between time of accountability and attitude ($F(1,56) = 4.40$, $p = .040$). A simple main effects analysis revealed that the attitude effect only approached significance when participants had been made accountable before the first decision ($F(1,56) = 3.30$, $p = .074$). In this case, participants with a pro-management attitude showed a lower variability of search across alternatives than participants with a pro-user attitude ($M = 0.48$ vs. $M = 0.88$), for participants who had been made accountable before the final decision this pattern was reversed ($M = 0.87$ vs. $M = 0.62$), but not significantly so. This relationship is illustrated in Figure 7.10.

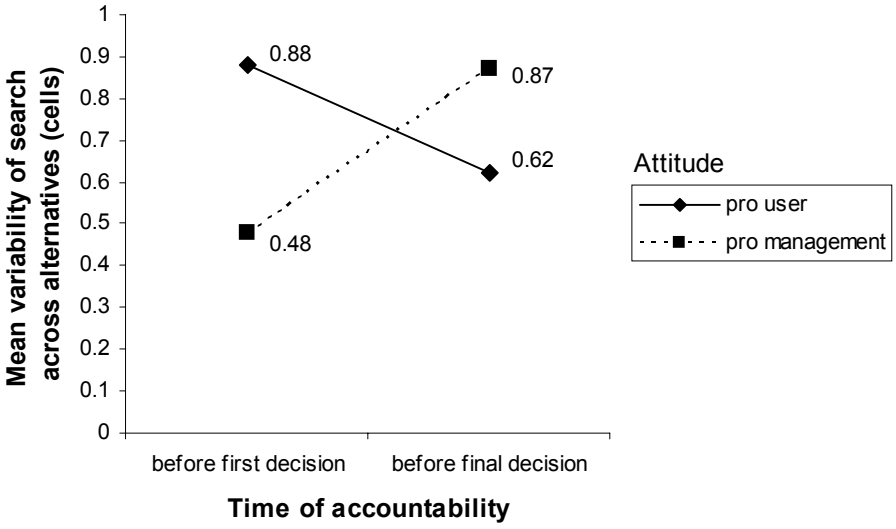


Figure 7.10. Mean variability of search across alternatives (based on different cells searched) during the final information search, as a function of time of accountability and attitude.

The ANOVA carried out on the variability of search across attributes calculated for different cells searched yielded a significant main effect of time of accountability ($F(1,56) = 4.15$, $p = .046$). As expected, participants who had been made accountable before the final decision showed a lower variability of search across attributes than participants who had been

made accountable before the first decision (M = 0.41 vs. M = 0.66). This effect was qualified, however, by a marginally significant interaction between time of accountability and type of accountability ($F(1,56) = 2.93$, $p = .092$). A simple main effects analysis indicated that time of accountability only had a significant effect on participants who had been made accountable to the users of the program ($F(1,56) = 7.03$, $p = .010$). Participants who had been made accountable to the users of the program before the first decision showed a higher variability of search across attributes than participants who had been made accountable only before the final decision (M = 0.73 vs. M = 0.25). For participants who had been made accountable to the management, this difference was not significant (M = 0.60 vs. M = 0.56, see Figure 7.11).

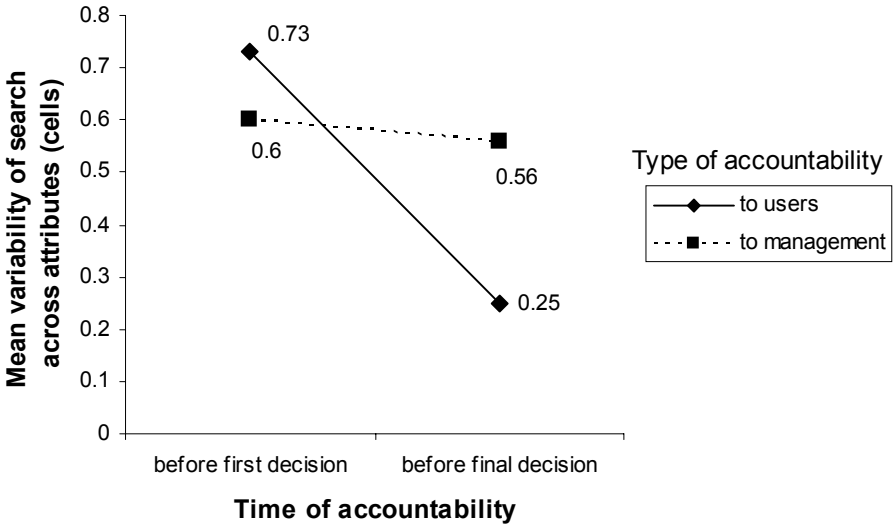


Figure 7.11. Mean variability of search across attributes (based on different cells searched) during the final information search, as a function of time of accountability and type of accountability.

4. Compensatory processing

Overall, information processing was significantly less compensatory than during the first search (0.52 vs. 0.60, $t(79) = 2.27$, $p = .026$, two-sided), which is consistent with a strategy of mainly searching information for the alternative one already has a preference for.

Although the predicted main effect of accountability on the extent of compensatory processing failed to reach standard significance levels ($F(1,76) = 2.34$, $p = .131$), it was in the expected direction. Accountable participants tended to process information in a more compensatory manner than non-accountable participants ($M = 0.55$ vs. $M = 0.40$). The main effect of attitude also showed a weak tendency to be significant ($F(1,74) = 2.43$, $p = .123$). Participants with a pro-user attitude tended to process information in a more compensatory manner than participants with a pro-management attitude ($M = 0.46$ vs. $M = 0.57$). This is opposite to what was found for the first search.

It had also been expected that participants who had been made accountable before the final decision would show a stronger tendency for compensatory processing than participants who had already been made accountable before the first decision, resulting in a significant main effect of time of accountability. This effect was not found. There was, however, a marginally significant interaction between time of accountability and attitude ($F(2,74) = 2.38$, $p = .099$). Simple main effects analyses suggested that the effect of time of accountability only tended to be significant for participants with a pro-user attitude ($F(2,74) = 2.40$, $p = .097$). Their information processing was significantly more compensatory when they had been made accountable before the final decision than when they had not been made accountable ($M = 0.60$ vs. $M = 0.28$, $t(79) = 2.11$, $p = .039$). It also tended to be more compensatory when they had been made accountable before the final decision than when they had been made accountable before the first decision ($M = 0.60$ vs. $M = 0.48$), but not significantly so. In addition, the analyses revealed that the effect of attitude tended to depend on the 'time of accountability' condition ($F(3,74) = 2.20$, $p = .095$). The only significant difference between participants with a pro-user attitude and participants with a pro-management attitude was obtained when they had been made accountable before the first decision. In this case, participants with a pro-management attitude showed a more compensatory information search than participants with a pro-user attitude ($M = 0.68$ vs. $M = 0.42$, $t(79) = 2.03$, $p = .046$).

The ANOVA testing the effects of type of accountability and attitude showed a weak tendency for a main effect of type of accountability ($F(2, 74) = 2.20$, $p = .118$). Post-hoc contrasts showed that participants who had been made accountable to the users of the program tended to have a more strongly compensatory information processing than participants who

had not been made accountable (M = 0.61 vs. M = 0.40, p = .161). The attitude main effect only showed a weak tendency to be significant (F(1,74) = 2.53, p = .116).

The results of the three-factorial ANOVAs testing the effect of time of accountability, type of accountability and attitude on the experimental groups only yielded the following results. Consistent with the results for depth of search and variability of search across alternatives, a significant interaction between time of accountability and attitude was found (F(1,56) = 4.34, p = .042). As before, simple main effects analyses suggested that participants' attitude only had a significant effect if they had been made accountable before the first rather than the final decision (F(1,56) = 4.48, p = .039). In this case, participants with a pro-management attitude processed information in a more compensatory manner than participants with a pro-user attitude (M = 0.68 vs. M = 0.42). For participants who had been made accountable before the final decision the pattern was reversed (M = 0.50 vs. M = 0.60), the difference was not significant, however. Figure 7.12 shows the extent of compensatory information processing, as a function of time of accountability and attitude.

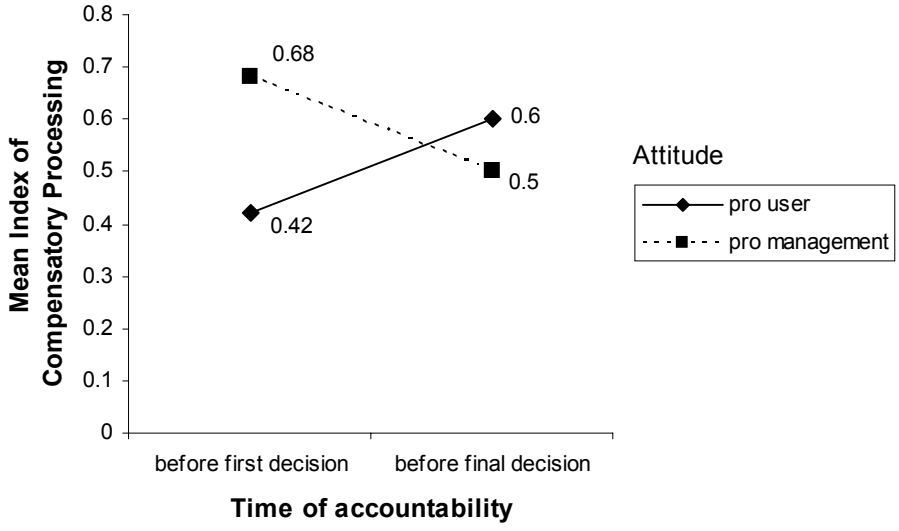


Figure 7.12. Mean Index of Compensatory Processing, as a function of time of accountability and attitude (a value of 1 indicates a maximum extent of compensatory processing).

5. Pattern of search (Payne Index)

The overall pattern of search was attribute-wise, as indicated by the negative value obtained (M = -0.41), but not significantly different to that of the first search (M = -0.34). This suggests that, overall, information processing tended to be non-compensatory.

It had been predicted that non-accountable participants would show a more attribute-wise search than accountable participants. The main effect of accountability in the 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro management) ANOVA employed to test this hypothesis was not significant, however. If anything, the observed means suggested the opposite; information search tended to be more attribute-wise for accountable than non-accountable participants (M = -0.43 vs. M = -0.29). The attitude main effect was not significant, nor was the interaction between accountability and attitude.

Another difference that had been predicted was that participants who had been made accountable before the first decision would show a more attribute-wise search than participants who had been made accountable before the final decision. This was not the case; in the ANOVA employed to test this hypothesis, the main effect of time of accountability was not significant, nor was the main effect of attitude or the interaction between these factors. The same was true for the analysis testing the effect of type of accountability.

Finally, no significant effects were obtained in the analysis testing the effects of time of accountability, type of accountability and attitude on the experimental conditions only.

6. Concentration of search

The preliminary conclusion that participants mainly searched information for the alternative they already had a preference for was also supported by the results for concentration of search on the chosen alternative. Concentration of search on the chosen alternative was higher during the final information search than during the first search, both for the total number of items searched and for different cells searched (M = 0.39 vs. M = 0.35 for total number of items searched, M = 0.33 vs. M = 0.22 for different cells searched). This

difference was only significant for the concentration of search for different cells searched, however ($t(79) = 2.94$, $p = .004$).

It had been predicted that accountable participants would concentrate their search less on the chosen alternative than participants who had not been made accountable. There was indeed a significant main effect of accountability for concentration of search on the chosen alternative calculated for the total number of items searched ($F(1,76) = 4.66$, $p = .034$). As expected, accountable participants focused their search less on the chosen alternative than non-accountable participants ($M = 0.35$ vs. $M = 0.57$). There was also a highly significant main effect of attitude ($F(1,76) = 11.56$, $p = .001$). Participants with a pro-user attitude concentrated their information search more on the chosen alternative than participants with a pro-management alternative ($M = 0.55$ vs. $M = 0.23$). The results for concentration of search of different cells were less clear-cut. The main effect of accountability only showed a weak tendency to be significant ($F(1,76) = 2.25$, $p = .138$). As before, accountable participants tended to concentrate their information search less on the chosen alternative than non-accountable participants ($M = 0.30$ vs. $M = 0.45$). The only significant effect was the attitude main effect ($F(1,76) = 6.79$, $p = .011$), which, again, suggested that participants with a pro-user attitude had a higher concentration of search on the chosen alternative than participants with a pro-management attitude ($M = 0.45$ vs. $M = 0.21$).

Time of accountability had a marginally significant effect on concentration of search for the total number of items searched ($F(2,74) = 2.43$, $p = .095$). Post-hoc simple contrasts indicated that participants who had been not been made accountable concentrated their search more on the chosen alternative than participants who had been made accountable before the final decision ($M = 0.57$ vs. $M = 0.32$, $p = .034$), and tended to concentrate their search more on the chosen alternative than participants who had been made accountable before the first decision ($M = 0.57$ vs. $M = 0.37$, $p = .080$, see Figure 7.13). Again, the main effect of attitude was significant ($F(1,74) = 14.48$, $p = .000$). For concentration of search for different cells searched, only the main effect of attitude was significant ($F(1,74) = 8.31$, $p = .005$). Participants with a pro-user attitude concentrated their search more on the chosen alternative than participants with a pro-management attitude ($M = 0.45$ vs. $M = 0.21$).

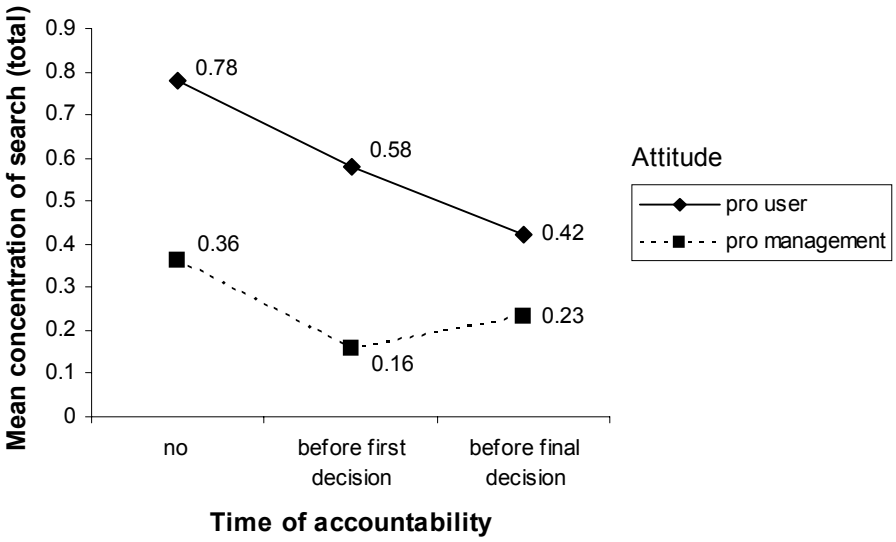


Figure 7.13. Mean concentration of the final information search on the chosen alternative (based on the total number of items searched), as a function of time of accountability and attitude.

Type of accountability was found to have a significant main effect on the concentration of search calculated for the total number of items searched ($F(2,74) = 3.40$, $p = .039$). Post-hoc simple contrasts showed that participants in the 'no accountability' condition showed a significantly higher concentration of search on the chosen alternative than participants in the 'accountability to users' conditions ($M = 0.57$ vs. $M = 0.28$, $p = .012$). Again, the main effect of attitude was highly significant ($F(1,74) = 15.06$, $p = .000$). There was no significant main effect of type of accountability on the concentration of search for different cells searched; only the attitude main effect was highly significant ($F(1,74) = 8.67$, $p = .004$).

Three-factorial ANOVAs testing the effects of time of accountability, type of accountability, and attitude on the concentration of search calculated for the total number of items searched as well as different cells searched, only yielded significant attitude main effects ($F(1,56) = 10.55$, $p = .002$ for total number of items searched, $F(1,56) = 7.08$, $p = .010$ for different cells searched). Participants with a pro-user attitude concentrated their search more on the chosen alternative than participants with a pro-management attitude ($M = 0.49$ vs. $M = 0.19$ for total number of items searched, $M = 0.41$ vs. $M = 0.19$ for different cells searched).

7. Biased search

It had been predicted that, after a first decision had been made, information search would be biased to bolster the previously chosen alternative, that is, participants would search information which they could expect to support their chosen alternative and devalue their non-chosen alternatives, resulting in a BSI that was significantly greater than 1 for the chosen alternative, significantly smaller than 1 for the average non-chosen alternative and significantly greater than 1 for the ratio of the chosen alternative to the average non-chosen alternative. This tendency was expected to be stronger for accountable than for non-accountable participants, being evident in a significant main effect of accountability. In addition, information search was expected to be particularly biased if participants had been made accountable before the final rather than the first decision, and if they experienced a value conflict because their own preferences opposed those of the group they had been made accountable to, resulting in significant interactions between time of accountability and attitude and type of accountability and attitude.

The overall mean BSI observed for the chosen alternative (M = 1.17) indicated that there was a bias in the sample to search more supporting than non-supporting information for the chosen alternative. The difference to a balanced value of 1 was highly significant ($t(79)$ = 3.32, p = .001). The same was true for the ratio of the BSI for the chosen to the BSI for the average non-chosen alternative (M = 1.26, $t(79)$ = 3.28, p = .002), suggesting that information search was biased to support the chosen over the average non-chosen alternative. The observed mean BSI for the non-chosen alternative (M = 1.03) was not significantly different from 1, however.

Two-factorial ANOVAs testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro-management) were performed on all dependent measures to test whether the bias in information search was stronger for accountable than for non-accountable participants. Contrary to expectations, no significant main effects of accountability were obtained. Accountable participants tended to show a larger BSI for the chosen alternative than non-accountable participants (M = 1.19 vs. M = 1.09), but the observed difference was not significant. For the average non-chosen alternative, a tendency in the opposite direction was

observed, although this was not significant either; accountable participants did not bias their information search against the average non-chosen alternative, whereas non-accountable participants tended to do so (M = 1.06 vs. M = 0.95). This may be seen to suggest that the expectation of having to justify their choice induced participants to pay more attention to information that contained possible counter-arguments. As a result, the ratio of the BSI for the chosen alternative to the BSI for the average non-chosen alternative hardly differed between conditions (accountability: M = 1.25, no accountability: M = 1.30). The ANOVA performed on this dependent measure yielded the only significant effect, a main effect of attitude ($F(1,76) = 4.67$, $p = .034$). Participants with a pro-user attitude showed a stronger bias to support their chosen over their average non-chosen alternative than participants with a pro-management attitude (M = 1.45 vs. M = 1.07).

In order to test whether the observed bias in information search was stronger for participants who had been made accountable before the final rather than the first decision, two factorial ANOVAs including the factors of time of accountability and attitude were conducted. The analysis performed on the BSI for the chosen alternative did not yield the expected main effect of time of accountability, but showed a marginally significant main effect of attitude ($F(1,74) = 3.17$, $p = .079$). Participants with a pro-user attitude tended to show a stronger bias to support their chosen alternative than participants with a pro-management attitude (M = 1.26 vs. M = 1.07). This effect was qualified by a marginally significant interaction between time of accountability and attitude ($F(2,74) = 2.34$, $p = .104$). Simple main effect analyses indicated that the strength of the attitude effect depended on when participants had been made accountable ($F(3,74) = 2.95$, $p = .038$). Participants with a pro-user attitude only showed a stronger bias to support their chosen alternative than participants with a pro-management attitude when they had been made accountable before the final decision (M = 1.43 vs. M = 0.98, $t(79) = 2.94$, $p = .004$). The analysis performed on the BSI for the average non-chosen alternative did not yield any significant effects, that for the ratio of the BSI for the chosen to the BSI for the average non-chosen alternative only yielded the already known main effect of attitude ($F(1,74) = 5.81$, $p = .018$).

The predictions concerning the effects of a value conflict on the degree of bias in participants' information search were tested with three-factorial ANOVAs including the

factors of time of accountability, type of accountability and attitude. These only included the experimental groups. The analysis performed on the BSI for the chosen alternative indicated a marginally significant main effect of attitude ($F(1, 56) = 3.43$, $p = .069$) and a marginally significant interaction between time of accountability and attitude ($F(1, 56) = 3.72$, $p = .059$). Participants with a pro-user attitude tended to bias their information search more strongly towards the chosen alternative than participants with a pro-management attitude ($M = 1.30$ vs. $M = 1.07$). However, this effect was only significant when they had been made accountable before the final decision ($M = 1.43$ vs. $M = 0.98$, $F(1, 56) = 7.21$, $p = .010$). There was no significant difference between participants with a pro-user and a pro-management attitude when they had been made accountable before the first decision ($M = 1.16$ vs. $M = 1.17$). This relationship is depicted in Figure 7.14.

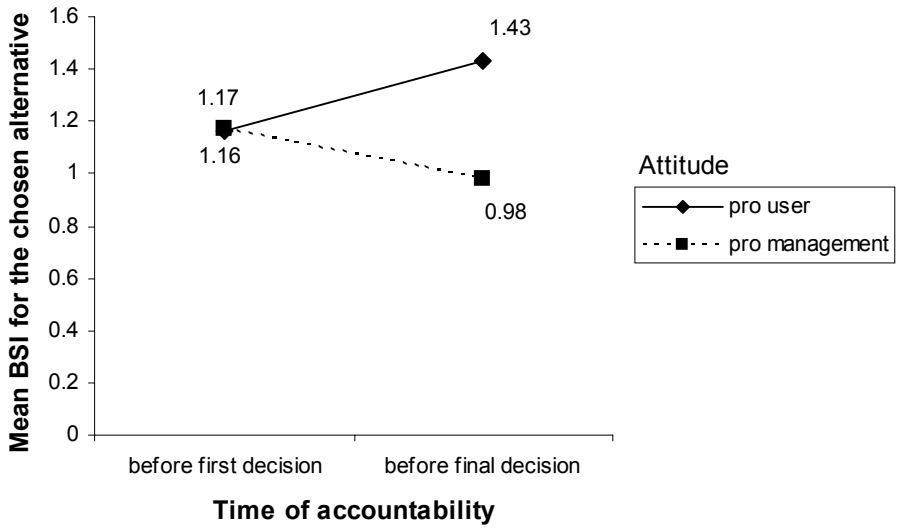


Figure 7.14. Mean BSI for the chosen alternative, as a function of time of accountability and attitude.

Similar results were obtained for the ratio of the BSI for the chosen alternative to the BSI for the average non-chosen alternative. The main effect of attitude was again marginally significant ($F(1,56) = 3.86$, $p = .055$, pro-user attitude: $M = 1.43$, pro-management attitude: $M = 1.08$). The interaction between time of accountability and attitude was not significant ($F(1,56) = 1.81$, $p = .183$), but simple main effects analyses indicated that, as already

suggested by the BSI for the chosen alternative, participants with a pro-user attitude biased their information search more strongly to support their chosen over their average non-chosen alternative than participants with a pro-management attitude when they had been made accountable before the final, but not when they had been made accountable before the first decision (accountability before the final decision: $M = 1.53$ vs. $M = 0.94$, $F(1,56) = 5.52$, $p = .022$, accountability before the first decision: $M = 1.33$ vs. $M = 1.22$). Although the triple interaction between time of accountability, type of accountability and attitude was not significant, an inspection of the observed means showed that participants with a pro-user attitude who had been made accountable to the management before the final decision exhibited a particularly strong bias to support their chosen over their average non-chosen alternative, suggesting that they experienced their value conflict as particularly strong (see Figure 7.15). No significant effects were obtained in the ANOVA analysing the BSI for the average non-chosen alternative.

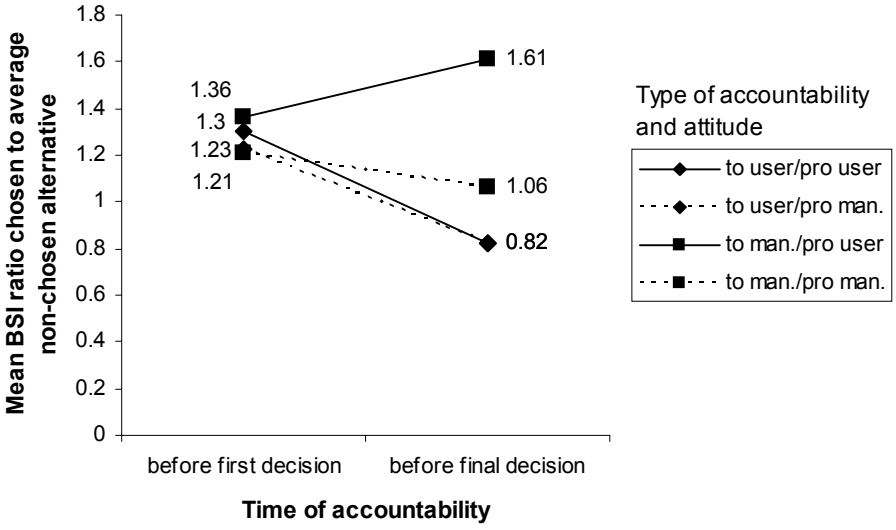


Figure 7.15. Mean ratio of the BSI for the chosen alternative to the BSI for the average non-chosen alternative.

Summary of the results for the final information search

The results for the final information search were consistent with predictions derived from the PMA, in that they suggested that accountability pressures tended to result in a deeper, less variable and more compensatory information search that was less strongly concentrated on the chosen alternative compared to no accountability. The difference between accountability before the final and the first decision was also significant in some cases, but tended to be smaller overall. Also, the final information search measures pointed to some of the predicted interactions between the experimental factors. As expected, a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in information processing that was geared towards protecting the chosen alternative if participants had not adjusted their choice to that assumed to be preferred by their audience. This was only true for participants with a pro-user attitude who had been made accountable to the management, however, because the general bias towards user considerations in the sample meant that participants with a pro-management attitude felt it easy to give in to user pressures. Because participants who had been made accountable before the first decision had partly adjusted to audience preferences already when making their first decision, as expected, this effect tended to be stronger for participants who had been made accountable before the final decision.

Overall, depth of the final information search was lower than that of the first. This was to be expected, given that participants had already developed a preference hierarchy, and that the information to be searched was largely redundant with that of the first information set and therefore should not have presented any surprises to the participants. As predicted, accountable participants searched more information than non-accountable participants. However, the hypothesis that participants who had been made accountable immediately prior to the final information search would search more information than participants who had already been made accountable before the first information search could only be confirmed for participants with a pro-user attitude. This was because, although participants with a pro-management attitude searched more information than participants with a pro-user attitude overall, and did so when they had either not been made accountable or when they had been made accountable before the first decision, they failed to do so when they had been made

accountable before the final decision, particularly when they had been made accountable to the management. This was true both for the total number of items searched and different cells searched. As a consequence of this effect, only the 'accountability to users' conditions showed a deeper information search than the 'no accountability' conditions. At this point, there is no readily apparent reason why participants with a pro-management attitude who had been made accountable to the management before the final decision searched comparatively little information. The analysis of participants' choices later on may shed some light on this. The significant main effect of attitude on depth of search is yet another indicator for the fact that participants with a pro-management attitude were concerned with both user and cost attributes, whereas participants with a pro-user attitude were primarily interested in user attributes.

The percentage of searched user information observed in the various conditions was consistent with a results pattern that would be expected if participants with a pro-user attitude who had been made accountable to the management were reluctant to adjust to these accountability pressures, whereas participants with a pro-management attitude who had been made accountable to the users were willing to. This had already been suggested by analyses investigating information evaluation processes after the previous decision and so receives further support. Participants with a pro-user attitude searched a higher percentage of user information when they had been made accountable to the management than when they had been made accountable to the users of the program, presumably because they had mostly chosen the user alternative before and needed to defend it against objections from a management-friendly audience. Participants with a pro-management attitude who had been made accountable to the users of the program, on the other hand, searched as much (and not less) user information as participants with a pro-user attitude, suggesting that they adjusted to the preferences of the audience they had been made accountable to. It had also been expected that participants who had been made accountable before the final decision would be unwilling to adjust to the preferences of their audience if these conflicted with their own, but that this would not be the case if participants had already been made accountable before the first decision. However, this effect could not be observed in the percentage of user information searched.

On the whole, the observed variability of search, as expected, tended to be lower for accountable than for non-accountable participants, but not significantly so. For variability of search across attributes, calculated for the total number of items searched, an opposite trend could be observed, suggesting that accountable participants, when they searched information repeatedly, tended to do so for some attributes but not for others to a larger extent than non-accountable participants, possibly to remind themselves of the advantages of the alternative they favoured. Variability of search across alternatives calculated for different cells searched was generally larger than variability of search across attributes. This indicates that participants only searched a subset of the available alternatives, but tended to search an equal number of attributes per alternative. The expected main effect of time of accountability on variability of search measures was only found for variability of search across attributes for different cells searched. As expected, it tended to be lower when participants had been made accountable before the final decision than when they had either not been made accountable or had been made accountable before the first decision. Variability of search across alternatives calculated for different cells searched was affected by time of accountability only when participants had a pro-user attitude. This was consistent with the results obtained for depth of search. For participants with a pro-user attitude, variability of search across alternatives was indeed lower when they had been made accountable before the final decision than when they had not been made accountable, whereas for participants with a pro-management attitude there was no significant difference, and they tended to show the lowest variability of search when they had already been made accountable before the first decision. The fact that participants' attitude moderated the effects of time of accountability was further supported by a significant interaction between these factors when the data for the experimental groups only were inspected. Significant effects were only obtained for variability of search measures related to the number of different cells searched. Where variability of search across alternatives was concerned, participants with a pro-user attitude showed a significantly larger variability of search than participants with a pro-management attitude when they had been made accountable before the first decision. For accountability before the final decision, the reverse tended to be true, but not significantly so. This is yet another indicator for a tendency described earlier, namely that participants with a pro-management attitude did not show their usual relatively balanced search of all alternatives when they had been made accountable

before the final decision, especially when they had been made accountable to the management. This meant that for variability of search across attributes, participants who had been made accountable to the management of the company did not show the expected smaller variability of search when they had been made accountable before the final compared to the first decision, whereas participants who had been made accountable to the users did.

The analyses performed on the Koele and Westenberg Index of Compensatory Processing, which combines measures of depth of search and variability of search across alternatives for different cells searched, not surprisingly, confirmed what had separately been found for these two measures. The information search of participants with a pro-management attitude was more compensatory than that of participants with a pro-user attitude when they had been made accountable before the first decision, whereas the opposite tended to be the case when they had been made accountable before the final decision. Also, participants with a pro-user attitude processed information in a more compensatory manner when they had been made accountable before the final decision than when they had not been made accountable. As expected, information processing tended to be more compensatory for accountable than non-accountable participants, but the observed difference was not significant. Overall, processing was also less compensatory than during the first information search. This effect was also evident in the observed pattern of search, which was generally attribute-wise and indicated a non-compensatory information search. The latter was, however, the only conclusion that could be drawn from the findings obtained regarding the pattern of search. The persistent failure to obtain any significant effects for this measure indicates problems with its sensitivity.

The findings for the extent of compensatory information processing were largely supported by those for the concentration of search measures. Concentration of search on the chosen alternative, as expected, was significantly lower for accountable than non-accountable participants when the total number of items searched was inspected, and tended to be lower for the number of different cells searched, suggesting that it was particularly repeated searches of information that focused on the chosen alternative. In most analyses, there was also a significant main effect of attitude, which indicated that participants with a pro-user attitude concentrated their information search more strongly on the chosen alternative than participants with a pro-management alternative. The analyses performed on the concentration of search

measure for the total number of items searched furthermore suggested that participants who had not been made accountable concentrated their information significantly more strongly on the chosen alternative than participants who had been made accountable immediately prior to this search, and marginally more strongly than participants who had already been made accountable before the first decision. Participants in the 'no accountability' conditions also showed a stronger concentration of search on the chosen alternative for the total number of items searched than participants in the 'accountability to users' conditions. In departure from the findings for depth and variability of search, the analyses carried out on the experimental groups only did not show any significant interactions, but only resulted in significant attitude main effects. In line with expectations, however, the highest concentration of search on the chosen alternative was observed for participants with a pro-user attitude who had been made accountable to the management before the first decision. They had experienced a value conflict, because they had not generally adjusted their choices to the preferences of their audience previously. This induced a need to concentrate on their preferred alternative, and as they did not experience an immediate pressure to prepare for a justification, they had the freedom to do so.

Finally, the analyses performed on the Biased Search Index (BSI), as a measure for confirmation bias, revealed that there was indeed a bias in the sample to search more supporting than non-supporting information for the chosen alternative, and to search information in such a way that the chosen alternative was supported over the average non-chosen alternative. Accountable participants, however, did not do so significantly more strongly than non-accountable participants. The analysis performed on the BSI for the average non-chosen alternative even suggested that accountable participants tended to search more information that supported the non-chosen alternatives than did participants who had not been made accountable. This may be seen as supporting the assumption that accountability raises interest in possible counter-arguments, and therefore increases attention to competing choices. The analyses performed on the BSI for the chosen alternative also revealed that participants with a pro-user attitude showed a stronger bias to support their chosen alternative than participants with a pro-management attitude, but only when they had been made accountable before the final decision. This was also found when the ratio of the BSI for the chosen

alternative to the BSI for the average non-chosen alternative was analysed. In this case, the effect was only significant, however, when the control group was not included in the analysis. When it was included, only the attitude main effect was significant. The nature of the observed interactions between time of accountability and attitude is consistent with other results that suggested that participants' information search was particularly non-balanced and biased towards the chosen alternative when they experienced a value conflict because they had been made accountable to a group whose preferences clashed with their own and they had not been willing to adjust to the preferences of their audience. This was particularly the case for participants with a pro-user attitude who had been made accountable to the management before the final decision. Hence, not surprisingly, the strongest bias to support the chosen over the average non-chosen alternative was observed in this condition.

Information evaluation and integration

1. Choices

It was expected that participants who had been made accountable before the first decision would be more likely to choose the user alternative if they had been made accountable to the users of the program, and more likely to choose a management alternative if they had been made accountable to the management of the company. Participants who had been made accountable only before the final decision, however (that is, after they had already made a commitment to one of the alternatives), were expected not to adjust their choice to the preference of the group they had been made accountable to but choose according to their own attitude instead. Participants who had not been made accountable were expected to show the same tendency. Table 7.9 shows the frequencies of user and management choices, as a function of time of accountability, type of accountability and attitude.

Table 7.9. Frequency of user and management choices at the final decision.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision				
			To Users		To Man.		To Users		To Man.		
Attitude	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Σ
User Alt.	6	5	8	6	3	4	7	3	6	4	52
Man. Alt	2	3	1	1	4	5	1	5	2	4	28
Σ	8	8	9	7	7	9	8	8	8	8	80

A Chi-squared analysis performed on the choices of participants who had been made accountable before the first decision revealed the expected effect; participants who had been made accountable to the users of the program chose the user alternative more often than a management alternative (14/16= 88%), whereas participants who had been made accountable to the management chose a management alternative more often than the user alternative (9/16 = 56%, χ^2 (1) = 6.79, p = .009). Participants who had only been made accountable before the final decision, on the other hand, did not adjust their choices to the group they were made accountable to (χ^2 (1) = 0.00, p = 1.000). As expected, their choices were a function of their own attitude instead. Participants with a pro-user attitude chose the user alternative in the majority of cases (13/16 = 81%), participants with a pro-management attitude, on the other hand, chose a management alternative more often than the user alternative (9/16 = 56%, χ^2 (1) = 4.80, p = .028). In the 'no-accountability' condition, however, although expected, this effect was not observed. Both participants with a pro-user and a pro-management attitude chose the user alternative more often than the management alternative (6/8 = 75% and 5/8 = 63%, respectively). The reason for this was that, surprisingly, three non-accountable participants who had been classified as having a pro-management attitude changed their initial decision from a management alternative to the user alternative. This had a considerable impact, given the small number of participants in this condition. Otherwise, changes of the initial decision were generally rare (10/80 = 12.5%) and exclusively in the direction of a change from a management alternative to the user alternative, again indicating the generally strong preference for user over cost attributes in the sample (see Table 7.10).

Table 7.10. Frequency of change of the initial decision.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision				
			To Users		To Man.		To Users		To Man.		
Attitude	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Pro User	Pro Man.	Σ
No Change	7	5	9	5	7	7	8	7	7	8	70
To User Alt	1	3	0	2	0	2	0	1	1	0	10
To Man. Alt	0	0	0	0	0	0	0	0	0	0	0
Σ	8	8	9	7	7	9	8	8	8	8	80

2. Attribute weights

Similar to the choice results, it was expected that participants who had been made accountable before the first decision would indicate adjustment of their attribute weights to the group they had been made accountable to, that is, participants who had been made accountable to the users would report higher user attribute weights than participants who had been made accountable to the management of the company. Participants who had been made accountable only before the final decision, on the other hand, would not show this adjustment, but instead report weights which corresponded with their attitude, resulting in an interaction between time of accountability and type of accountability and time of accountability and attitude. The mean sum of user attribute weights observed before and after making the final decision as well as their change from the second to the third elicitation is shown in Table 7.11.

Three-factorial ANOVAs testing the effects of time of accountability (before the first decision vs. before the final decision), type of accountability (to users of the program vs. to the management of the company) and attitude (pro-user vs. pro-management) were conducted to test the above hypotheses. The control group was not included in the analyses.

The ANOVA performed on the mean sum of normalised weights for user attributes contained in the *first* information set after the final decision revealed a significant main effect of attitude ($F(1,56) = 4.28$, $p = .043$). For participants with a pro-user attitude, a higher sum of user attribute weights was obtained than for participants with a pro-management attitude (M =

0.71 vs. M = 0.64). The analysis also revealed the expected interaction between time of accountability and type of accountability (F(1,56) = 9.09, p = .004). A simple main effects analysis showed that when participants had been made accountable before the first decision, the mean sum of user attribute weights was significantly higher in the 'accountability to users' condition than in the 'accountability to the management' condition (M = 0.72 vs. M = 0.62, F(1,56) = 8.37, p = .005). For participants who had been made accountable before the final decision, this difference was not significant (M = 0.69 vs. M = 0.65, see Figure 7.16). This was in line with predictions. Also, when participants had been made accountable to the users of the program, their sum of user attribute weights was significantly higher when this had happened before the first decision than when it had happened before the final decision (M = 0.72 vs. M = 0.65, F(1,56) = 4.24, p = .044). When participants had been made accountable to the management of the company, however, the opposite was true, the sum of user attribute weights was higher when participants had only just been made accountable to the management than when they had already been made accountable before the first decision (M = 0.62 vs. M = 0.69, F(1,56) = 4.87, p = .032). This makes sense, if one bears in mind that many of those participants had initially chosen the user alternative (9/16 = 56%). Subsequent accountability to the management created a strong pressure to either adjust to the preferences of the group they had been made accountable to or bolster their decision. All participants seem to have chosen a bolstering strategy, by selecting the user alternative again and stressing the importance of user attributes on which their chosen alternative fared particularly well.

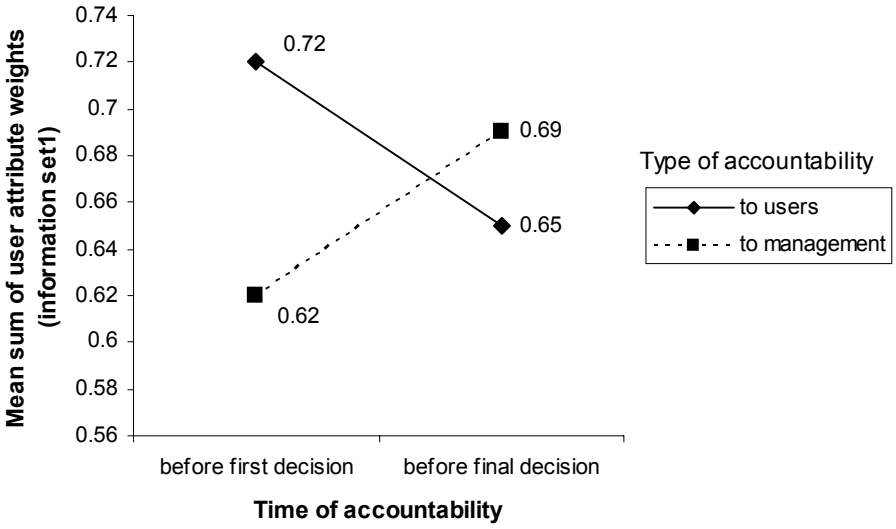


Figure 7.16. Mean sum of user attribute weights after the final decision (information set 1), as a function of time of accountability and type of accountability.

Unexpectedly, the ANOVA performed on the sum of weights for the parallel user attributes contained in the *second* information set only yielded a significant main effect of attitude ($F(1,55) = 5.38$, $p = .024$). Again, for participants with a pro-user attitude the sum of user attribute weights was higher than for participants with a pro-management attitude (M = 0.72 vs. M = 0.64). For the sum of weights based on the user attributes contained in *both* information sets, the time of accountability by type of accountability interaction remained significant, however ($F(1,55) = 4.64$, $p = .036$). Post-hoc analyses showed a similar pattern as before; for participants who had been made accountable before the first decision the observed mean sum of user attribute weights tended to be higher when they had been made accountable to the users of the program than the management of the company (M = 0.71 vs. M = 0.64, $F(1,55) = 3.51$, $p = .066$), whereas this difference was not significant when participants had only been made accountable before the final decision (M = 0.66 vs. M = 0.70).

Table 7.11. Mean sum of user attribute weights after the first and final decision, as a function of time of accountability, type of accountability and attitude.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
sum of user attribute weights										
set 1 attributes										
after final decision	.775	.631	.733	.705	.630	.606	.670	.625	.748	.642
after first decision	.788	.668	.744	.743	.609	.622	.664	.639	.743	.620
difference	-0.013	-0.037	-0.111	-0.038	-0.021	-0.160	0.006	-0.140	0.005	0.022
set 2 attributes										
after final decision	.757	.643	.744	.663	.696	.652	.694	.647	.760	.637
after first decision	.713	.628	.753	.694	.738	.655	.735	.629	.778	.630
difference	0.044	0.015	-0.009	-0.031	-0.042	-0.002	-0.041	0.019	-0.018	0.014*
set 1 + 2 attributes										
after final decision	.763	.637	.739	.673	.652	.631	.680	.635	.754	.643
after first decision	.750	.654	.748	.712	.661	.637	.696	.631	.757	.625
difference	0.013	-0.017	-0.008	-0.038	-0.009	-0.006	-0.016	0.004	-0.003	0.017*

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.
* Because of missing values for the third elicitation, the data for one participant had to be excluded from the analysis, resulting in inconsistent difference values.

The analyses performed on the *change* in the sum of user attribute weights from after the first to after the final decision did not yield any significant effects, neither for the attributes from the first information set nor for those from the second information set or both combined. The differences obtained were mainly negative, indicating regression rather than differentiation and consolidation effects.

3. Overall evaluation of alternatives

The dependent variables were the same as previously analysed, except that the differences between MAUT predictions were calculated subtracting the predictions obtained after the first decision from those obtained after the final decision. The results are presented in Table 7.12 and Table 7.13.

3a. Rating chosen - average non-chosen alternative

It had been predicted that, for participants who had been made accountable in any way, the change in the judgmental difference between their chosen and average non-chosen alternative from the first to the final decision as well as the difference observed after the final decision would be smaller than for participants who had not been made accountable. In order to compare the experimental conditions to the 'no accountability' conditions, 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVAs were performed. For the *change* in the judgmental difference between the chosen and average non-chosen alternative, the expected main effect of accountability was not found, but a marginally significant interaction between accountability and attitude was obtained ($F(1,76) = 3.12$, $p = .082$). Simple main effects analyses indicated that when participants had not been made accountable, those with a pro-management attitude increased the judgmental difference between their chosen and average non-chosen alternative, whereas participants with a pro-user attitude decreased it ($M = 6.46$ vs. $M = -6.25$, $t(79)= 1.79$, $p = .077$). When participants had been made accountable, on the other hand, there was a decrease in both groups ($M = -1.08$ vs. $M = -2.34$). In the analysis of the judgmental difference between the chosen and average non-chosen alternative after the final decision, the expected main effect of accountability

Table 7.12. Measures of the evaluative difference between the chosen and average non-chosen alternative after the first and final decision.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
1. rating chosen - average non-chosen alternative										
after final decision	41.25	40.63	35.00	36.67	31.19	28.96	34.58	29.17	51.79	29.58
after first decision	47.50	34.17	35.04	41.19	35.24	30.44	43.83	27.92	43.25	34.58
difference	-6.25	6.46	-0.04	-4.52	-4.05	-1.48	-9.25	1.25	8.54	-5.00
2. MAUT prediction chosen - average non-chosen alternative										
set 1 attributes										
after final decision	32.21	12.54	23.53	33.73	33.89	8.88	21.06	24.47	31.98	22.12
after first decision	34.21	16.28	34.64	26.74	35.86	12.80	27.76	25.52	28.52	21.57
difference	-2.00	-3.74	-11.11	6.99	-1.95	-3.91	-6.71	-1.05	3.46	0.55
set 2 attributes										
after final decision	35.32	30.09	38.61	25.82	23.44	21.94	28.09	21.54	40.60	10.37
after first decision	17.96	7.76	39.27	14.39	18.00	6.87	33.67	9.89	29.76	14.06
difference	17.36	22.33	-0.66	11.43	11.02*	15.07	-5.58	11.65	10.84	-2.34*
set 1 + 2 attributes										
after final decision	31.88	18.75	31.35	27.90	31.39	15.20	23.50	23.02	35.13	14.44
after first decision	31.28	16.30	40.85	22.54	24.42	9.98	32.73	18.61	28.81	22.06
difference	0.61	2.45	-9.50	5.36	9.65*	5.22	-9.23	4.41	6.32	-7.69*

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.
*Because of missing values for the third elicitation, the data for one participant had to be excluded from the analysis, resulting in inconsistent difference values.

Table 7.13. Measures of the evaluative difference between the user and average management alternative after the first and final decision.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
3. rating user - average management alternative										
after final decision	37.92	24.79	30.56	36.67	27.38	20.07	29.58	10.83	39.29	22.92
after first decision	32.50	19.17	29.11	25.95	18.10	2.30	33.83	7.92	29.92	18.75
difference	5.42	5.63	1.44	10.71	9.29	17.78	-4.25	2.92	9.38	4.17
4. MAUT prediction user - average man. Alternative										
set 1 attributes										
after final decision	41.66	7.98	26.16	26.59	29.29	12.12	23.23	5.18	37.84	23.78
after first decision	38.03	11.38	35.67	21.05	36.00	9.51	30.17	7.70	31.46	21.38
difference	3.63	-3.40	-9.50	5.54	-5.09*	2.61	-6.94	-2.52	6.39	2.39
set 2 attributes										
after final decision	47.87	23.39	40.67	30.48	63.88	31.66	31.43	18.70	49.91	19.78
after first decision	32.72	14.70	40.72	24.65	40.92	21.96	37.59	14.32	37.97	11.87
difference	15.15	8.68	-0.05	5.83	24.72*	9.70	-6.15	4.38	11.94	9.58*
set 1 + 2 attributes										
after final decision	42.51	13.22	33.84	26.23	42.65	21.08	26.37	10.38	42.95	20.18
after first decision	38.85	17.06	41.98	25.57	34.80	18.37	35.78	15.99	35.69	21.56
difference	3.66	-3.84	-8.14	0.66	-4.29*	2.71	-9.41	-5.62	7.26	-1.38*

Note: Discrepancies between the differences between the values obtained at the first and second elicitation and the given difference values are due to rounding errors.
* Because of missing values for the third elicitation, the data for one participant had to be excluded from the analysis, resulting in inconsistent difference values.

was marginally significant ($F(1,76) = 2.99$, $p = .088$). In line with predictions, participants who had been made accountable tended to show a smaller judgmental difference between their chosen and their average non-chosen alternative than participants who had been made accountable (M = 34.56 vs. M = 40.94). This lends further support to the conclusion drawn from previous results, namely that accountability seems to diminish the overt judgmental difference between the chosen and the average non-chosen alternative.

It was furthermore expected that participants who experienced conflict because their own preferences contradicted those of the group they had been made accountable to, would show a smaller judgmental difference between their chosen and their average non-chosen alternative than participants who did not experience this conflict. This was assumed to be more likely when participants had been made accountable before the final rather than the first decision, resulting in significant interactions between type of accountability and attitude and time of accountability and attitude. In order to test this hypothesis, 2 (time of accountability: before the first decision vs. before the final decision) by 2 (type of accountability: to users vs. to the management) by 2 (attitude: pro-user vs. pro-management) factorial ANOVAs were conducted. The control group was not included in the analyses.

The analysis performed on the *change* in the judgmental difference between the chosen and average non-chosen alternative from the first to the final decision revealed a significant interaction between all three experimental variables ($F(1,56) = 4.84$, $p = .032$). Simple main effects analyses indicated that the interaction between time of accountability and type of accountability was only significant for participants with a pro-user attitude ($F(1,56) = 4.76$, $p = .033$), and that the interaction between type of accountability and attitude was only significant for participants who had been made accountable before the final decision ($F(1,56) = 5.84$, $p = .019$). When participants had a pro-user attitude, they showed a larger decrease in the judgmental difference between their chosen and average non-chosen alternative when they had been made accountable to the management rather than the users before the first decision (M = -4.05 vs. M = -0.04), but when they had been made accountable before the final decision, they showed an increase in the judgmental difference between their chosen and average non-chosen alternative when they had been made accountable to the management, and a decrease when they had been made accountable to the users (M = 8.54 vs. M = -9.25). Participants with a pro-management

attitude tended to show the opposite pattern, but not significantly so; they showed a larger decrease in the judgmental difference between their chosen and average non-chosen alternative when they had been made accountable to the users rather than the management before the first decision (M = -4.52 vs. M = -1.48), but when they had been made accountable before the final decision, they showed an increase in the judgmental difference between their chosen and average non-chosen alternative when they had been made accountable to the users, and a decrease when they had been made accountable to the management (M = 1.25 vs. M = -5.00).

Also, when participants had been made accountable before the final decision, those with a pro-management attitude slightly increased the judgmental difference between their chosen and average non-chosen alternative when they had been made accountable to the users of the program, whereas those with a pro-user attitude decreased it (M = 1.25 vs. M = -9.25). For participants who had been made accountable to the management, the opposite was observed; those with a pro-user attitude increased the judgmental difference between their chosen and average non-chosen alternative, whereas those with a pro-management attitude decreased it (M = 8.54 vs. M = -5.00). When participants had been made accountable before the first decision, on the other hand, those with a pro-management attitude decreased the judgmental difference between their chosen and average non-chosen alternative more than those with a pro-user attitude when they had been made accountable to the users of the program (M = -4.52 vs. M = -0.04), but when participants had been made accountable to the management of the company, those with a pro-management attitude decreased the judgmental difference between their chosen and average non-chosen alternative less than those with a pro-user attitude (M = -1.48 vs. M = -4.05). The latter interaction was not significant, though. These relationships are depicted in Figure 7.17.

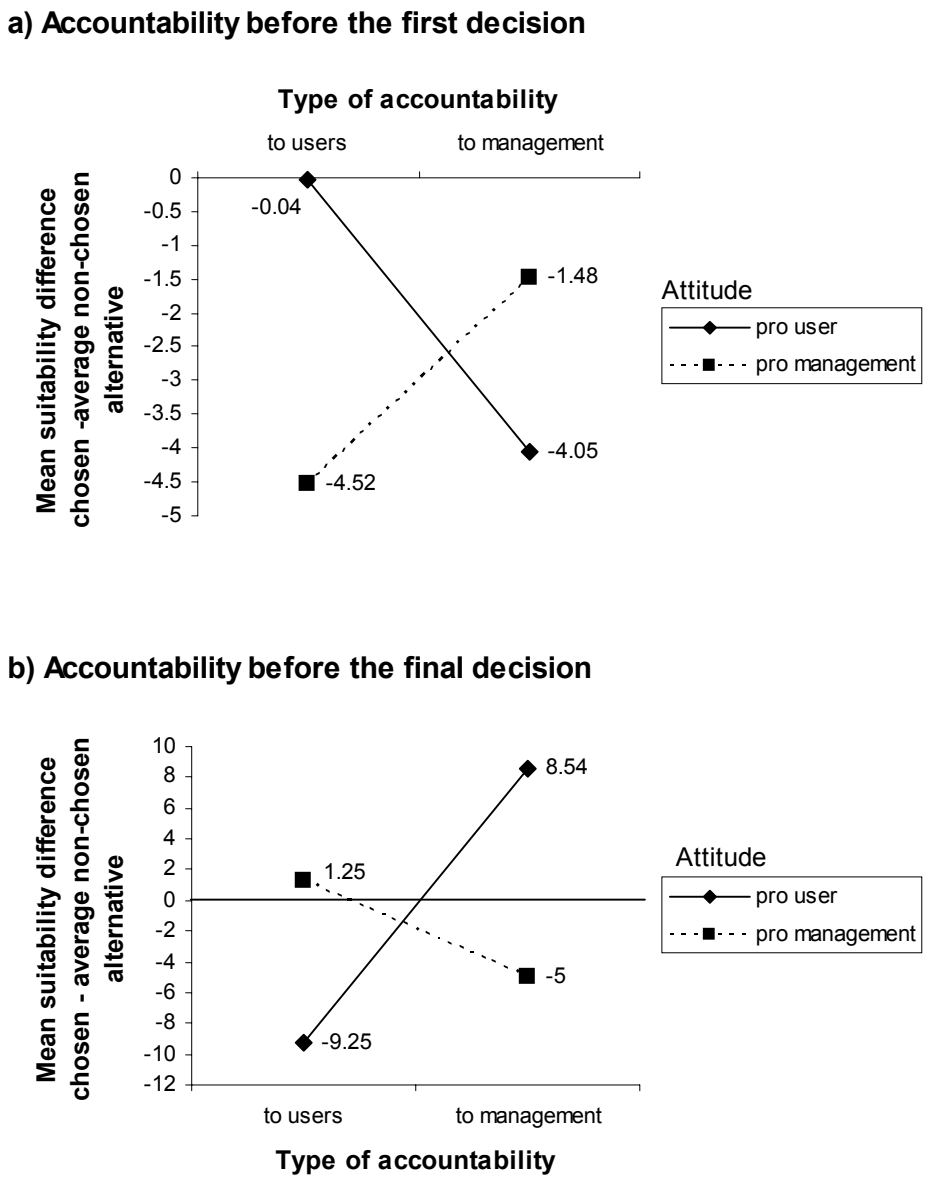


Figure 7.17. Mean change of the difference in overall suitability between the chosen and average non-chosen alternative from the first to the final decision.

This pattern of results suggests an interesting regularity; a conflict between participants' own preferences and the preferences of the audience they had been made accountable to resulted in *decreased* consolidation when participants had already been made accountable before the first decision, but resulted in an *increased* consolidation when they had only been made accountable before the final decision. This is consistent with the assumption that participants who had been made accountable before the first decision had adjusted to the preferences of their audience, and

therefore did not need to increase the evaluative difference between their chosen and average non-chosen alternative any further, whereas participants who had only been made accountable before the final decision had not, and therefore needed to further increase the evaluative difference between their chosen and average non-chosen alternative.

The analysis performed on the judgmental difference between the chosen and average non-chosen alternative obtained after the final decision yielded a significant main effect of attitude ($F(1,56) = 5.65$, $p = .021$), a significant interaction between time of accountability and attitude ($F(1,56) = 5.21$, $p = .026$) and a significant interaction between time of accountability and type of accountability ($F(1,56) = 6.04$, $p = .017$). The expected interaction between type of accountability and attitude was only marginally significant ($F(1,56) = 3.05$, $p = .086$). For participants with a pro-user attitude the judgmental difference between the chosen alternative and the average non-chosen alternative was higher than for participants with a pro-management attitude ($M = 38.26$ vs. $M = 30.85$). This effect was qualified, however, by the significant interaction between time of accountability and attitude and the marginally significant interaction between type of accountability and attitude. Simple main effect analyses indicated that if participants had been made accountable before the final decision, those with a pro-user attitude showed a larger judgmental difference between their chosen and their average non-chosen alternative than participants with a pro-management attitude ($M = 43.19$ vs. $M = 29.38$, $F(1,56) = 10.95$, $p = .002$). This difference was not significant if they had been made accountable before the first decision ($M = 33.33$ vs. $M = 32.33$). Also, participants with a pro-user attitude showed a significantly larger judgmental difference between their chosen and average non-chosen alternative when they had been made accountable before the final, compared to the first decision ($M = 43.19$ vs. $M = 33.33$, $F(1,56) = 5.80$, $p = .019$). For participants with a pro-management attitude, this was not the case ($M = 29.38$ vs. $M = 32.33$).

The simple main effects analyses employed to explore the interaction between type of accountability and attitude indicated that when participants had been made accountable to the management of the company, those with a pro-user attitude showed a significant larger difference between the suitability judgement for their chosen and average non-chosen alternative than those with a management attitude ($M = 42.18$ vs. $M = 29.25$, $F(1,56) = 8.50$, $p = .005$). There was no significant difference between participants with a pro-user attitude and with a pro-management

attitude when participants had been made accountable to the users of the program (M = 34.80 vs. M = 32.67).

Finally, simple main effect analyses suggested that participants who had been made accountable to the management showed a significantly larger judgmental difference between the their chosen and average non-chosen alternative if this had happened before the final rather than the first decision (M = 40.69 vs. M = 29.94, $F(1,56) = 6.41$, $p = .014$), whereas there was no significant difference when they had been made accountable to the users of the program (M = 31.88 vs. M = 35.73). Also, if participants had been made accountable before the final decision, those accountable to the users of the program showed a smaller judgmental difference than those accountable to the management (M = 31.88 vs. M = 40.69, $F(1,56) = 4.46$, $p = .039$). The opposite tended to be true if accountability had been introduced before the first decision, but not significantly so (M = 35.73 vs. M = 29.94).

This pattern of results seems to suggest that participants showed a *larger*, not smaller, judgmental difference between their chosen and their average non-chosen alternative after the final decision as a result of conflict between their own preferences and the preferences of the group they had been made accountable to. Also, this seems to only have been the case when the conflict was between having a pro-user attitude and having been made accountable to the management. Participants with a pro-user attitude who had been made accountable to the management showed a significantly higher difference between the suitability judgement for their chosen and average non-chosen alternative than participants with a pro-management attitude who had been made accountable to the users of the program (M = 42.18 vs. M = 32.67, $t(28) = 1.79$, p =.043, one-sided). Having a pro-management attitude when being faced with accountability to the users of the program does not seem to have created any strong conflict.

3b. MAUT prediction chosen - average non-chosen alternative

Because it was a covert measure, it had been expected that participants who had been made accountable would increase the difference between the MAUT prediction of the overall evaluation of their chosen and average non-chosen alternative more strongly than participants

who had not been made accountable. It had also been predicted that this difference would be significantly larger for accountable than non-accountable participants after the final decision. The 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVAs employed to test these hypotheses did not reveal any significant effects for the *change* in MAUT predictions based on the attributes contained in the *first* information set. A main effect of accountability was found in the analysis of the change in the MAUT predicted evaluative difference between the chosen and average non-chosen alternative based on the attributes from the *second* information set ($F(1,74) = 6.61$, $p = .012$). However, contrary to expectations, participants who had been made accountable showed a smaller increase in the evaluative difference between their chosen and average non-chosen alternative than participants who had not been made accountable ($M = 6.37$ vs. $M = 19.84$). No other effects were significant. Finally, the analysis performed on the change in the difference in MAUT predictions between the chosen and average non-chosen alternative combining the evaluation of attribute outcomes and weights from *both* information sets did not reveal any significant effects.

The analysis performed on the MAUT predicted difference between the chosen and average non-chosen alternative *after the final decision* only revealed a significant main effect of attitude when the prediction was based on the attributes from the *first* information set ($F(1,75) = 6.86$, $p = .011$). Again, participants with a pro-user attitude showed a larger difference between the MAUT prediction for their chosen and average non-chosen alternative than participants with a pro-management attitude ($M = 28.13$ vs. $M = 19.73$). Although the interaction between accountability and attitude failed to reach standard significance levels ($F(1,75) = 2.15$, $p = .147$), simple main effects analyses indicated that the effect of attitude depended on the accountability condition ($F(2,75) = 3.44$, $p = .037$). The difference between participants with a pro-user and a pro-management attitude was only significant in the 'no accountability' conditions ($M = 32.21$ vs. $M = 12.54$, $t(79) = 2.29$, $p = .025$), but not in conditions where participants had been made accountable ($M = 27.08$ vs. $M = 21.53$). No significant effects were obtained in the analysis of the predictions based on the attributes from the *second* information set. There was again a weak tendency for participants with a pro-user attitude to show a larger difference between the MAUT prediction for their chosen and average non-chosen alternative than participants with a pro-management attitude ($M = 33.85$ vs. $M = 22.15$), but it was not significant ($F(1,74) = 2.13$, $p = .149$). For the prediction based on the combined attributes from *both* information sets only the

already known main effect of attitude was found ($F(1,73) = 6.82$, $p = .011$). The difference between the MAUT prediction for the chosen and the average non-chosen alternative was again larger for participants with a pro-user attitude than for participants with a pro-management attitude ($M = 30.61$ vs. $M = 19.68$).

Three-factorial ANOVAs including the 'type of accountability' factor, the 'time of accountability' factor' and the 'attitude' factor were employed to test the hypothesis that participants who experienced a conflict, because their own preferences contradicted those of the group they had been made accountable to, would show a larger evaluative difference between their chosen and their average non-chosen alternative than participants who did not experience this conflict, and that this effect would be stronger for participants made accountable before the final rather than the first decision, resulting in significant interactions between type of accountability and attitude and time of accountability and attitude.

The analysis performed on the *change* in the difference between the MAUT predictions for the chosen and the average non-chosen alternative, based on the attributes contained in the *first* information set, showed indeed a marginally significant interaction between type of accountability and attitude ($F(1,55) = 3.18$, $p = .080$). Simple main effects analyses indicated that the effect of attitude was only significant when participants had been made accountable to the users of the program ($F(1,55) = 4.49$, $p = .039$). In this case, participants with a pro-user attitude decreased the MAUT predicted evaluative difference between their chosen and average non-chosen alternative, whereas participants with a pro-management attitude increased it ($M = -8.91$ vs. $M = 2.97$). When participants had been made accountable to the management, the opposite tendency could be observed, but was not significant ($M = 0.76$ vs. $M = -1.68$).

The ANOVA carried out on the change in the evaluative difference predicted from the evaluation of parallel attributes from the *second* information set also showed a marginally significant interaction between type of accountability and attitude ($F(1,54) = 3.97$, $p = .051$). Simple main effects analyses suggested that the effect of type of accountability was only significant for participants with a pro-user attitude ($F(1,54) = 4.22$, $p = .045$). In this case, participants who had been made accountable to the management of the company showed an increase in the MAUT predicted evaluative difference between their chosen and average non-

chosen alternative, whereas participants who had been made accountable to the users of the program showed a decrease (M = 10.93 vs. M = -3.12). Participants with a pro-management attitude, on the other hand, showed an increase both when they had been made accountable to the management and to the users (M = 7.45 vs. M = 11.54). Also, the effect of attitude was only significant, when participants had been made accountable to the users of the program (F(1,54) = 4.81, p = .033). In the 'accountability to users' conditions, participants with a pro-management attitude increased the MAUT predicted evaluative difference between their chosen and average non-chosen alternative, whereas participants with a pro-user attitude decreased it (M = 11.54 vs. M = -3.12). For participants who had been made accountable to the management, an increase was observed in both conditions (M = 6.36 vs. M = 10.93).

Not surprisingly, given that the analyses performed on the predictions for the attributes from the first and second information set separately had already shown interaction effects between type of accountability and attitude, this effect was also significant in the analysis of the predictions based on the attributes contained in *both* information sets (F (1,53) = 9.44, p = .003). As before, the effect of type of accountability was only significant for participants with a pro-user attitude (F(1,53) = 9.96, p = .003), and the effect of attitude was only significant when participants had been made accountable to the users of the program (F(1,53) = 7.48, p = .008). When participants had a pro-user attitude, those who had been made accountable to the management of the company showed an increase in the MAUT predicted evaluative difference between their chosen and average non-chosen alternative, whereas participants who had been made accountable to the users of the program showed a decrease (M = 7.60 vs. M = -9.36). The opposite pattern was observed when participants had a pro-management attitude; they showed a decrease, when they had been made accountable to the management and an increase when they had been made accountable to the users, although this effect was not significant (M = 0.43 vs. M = 4.85). Also, when participants had been made accountable to the users of the program, those with a pro-management attitude increased the MAUT predicted evaluative difference between their chosen and average non-chosen alternative, whereas participants with a pro-user attitude decreased it (M = 4.85 vs. M = -9.37). For participants who had been made accountable to the management, on the other hand, the effect was reversed, but not significant (M = 0.43 vs. M = 7.60).

The analyses performed on the MAUT predicted evaluative differences between the chosen and average non-chosen alternative obtained *after the final decision* yielded the following results. The ANOVA carried out on the predictions based on the attributes from the first information set yet again yielded a significant interaction between type of accountability and attitude ($F(1,55) = 8.77$, $p = .005$). Simple main effects analyses indicated that the effect of attitude was only significant when participants had been made accountable to the management of the company ($F(1,55) = 8.87$, $p = .004$). In this case, participants with a pro-user attitude showed a larger difference between the MAUT prediction for their chosen and average non-chosen alternative than participants with a pro-management attitude ($M = 32.80$ vs. $M = 15.11$). When participants had been made accountable to the users of the program, a reverse set of means could be observed ($M = 22.36$ vs. $M = 28.79$), although the difference between them was not significant. Also, the effect of type of accountability was only significant for participants with a pro-management attitude ($F(1,55) = 5.65$, $p = .021$). When participants had a pro-management attitude, those who had been made accountable to the users of the program showed a larger MAUT predicted evaluative difference than participants who had been made accountable to the management of the company ($M = 28.79$ vs. $M = 15.11$), whereas this effect was reversed when participants had a pro-user attitude, although only marginally significantly so ($M = 22.36$ vs. $M = 32.80$, $F(1,55) = 3.31$, $p = .074$). Although the three-way interaction between time of accountability, type of accountability and attitude failed to reach standard significance levels ($F(1,54) = 1.80$, $p = .186$), simple main effects analyses indicated that the interaction between type of accountability and attitude was only significant for participants who had been made accountable before the first decision ($F(1,55) = 8.97$, $p = .004$). Also, the interaction between time of accountability and type of accountability was marginally significant for participants with a pro-management attitude ($F(1,55) = 3.87$, $p = .054$), but not for participants with a pro-user attitude. These relationships are shown in Figure 7.18.

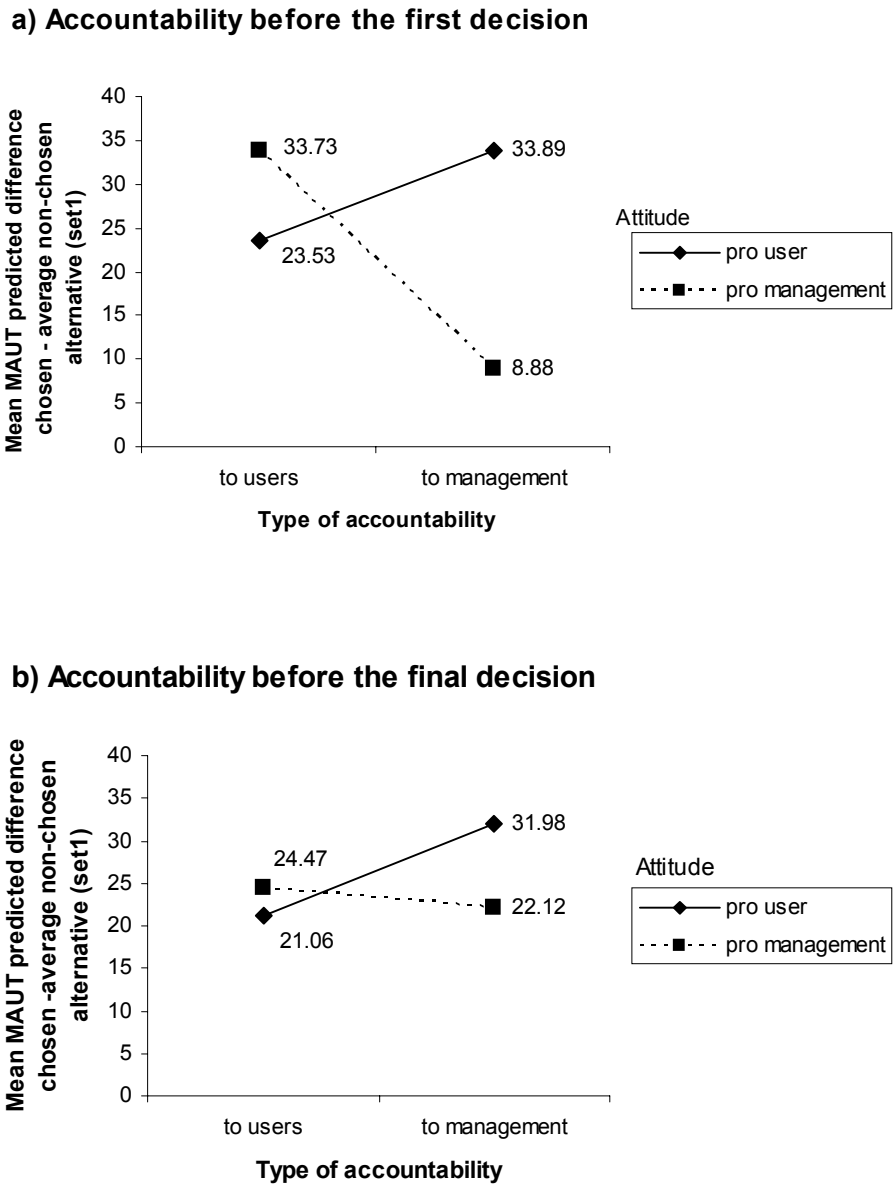


Figure 7.18. Mean MAUT predicted evaluative difference between the chosen and average non-chosen alternative after the final decision (based on the attributes from information set 1).

The ANOVA performed on the difference between the MAUT predictions for the chosen and the average non-chosen alternative based on the attributes contained in the *second* information set somewhat unexpectedly only yielded a significant main effect of attitude ($F(1,54) = 4.23$, $p = .044$). For participants with a pro-user attitude the obtained difference was larger than for participants with a pro-management attitude ($M = 33.47$ vs. $M = 20.10$).

Finally, the predictions based on the combined attributes from *both* information sets showed a significant main effect of attitude ($F(1,53) = 5.60$, $p = .022$) as well as a marginally significant interaction between type of accountability and attitude ($F(1,53) = 3.65$, $p = .061$). For participants with a pro-user attitude the obtained difference was larger than for participants with a pro-management attitude ($M = 30.27$ vs. $M = 19.91$). Simple main effects analyses suggested, however, that the effect of attitude was significant when participants had been made accountable to the management of the company ($M = 33.69$ vs. $M = 14.87$, $F(1, 53) = 8.56$, $p = .005$), but not when they had been made accountable to the users of the program ($M = 27.66$ vs. $M = 25.30$). Also, there was a marginally significant effect of type of accountability for participants with a pro-management attitude only ($F(1,53) = 3.16$, $p = .081$). In this case, participants who had been made accountable to the users of the program tended to show a larger difference between the MAUT predictions for their chosen and average non-chosen alternative than participants who had been made accountable to the management ($M = 25.30$ vs. $M = 14.87$). For participants with a pro-user attitude, the difference was not significant ($M = 27.66$ vs. $M = 33.69$).

3c. Rating user - average management alternative

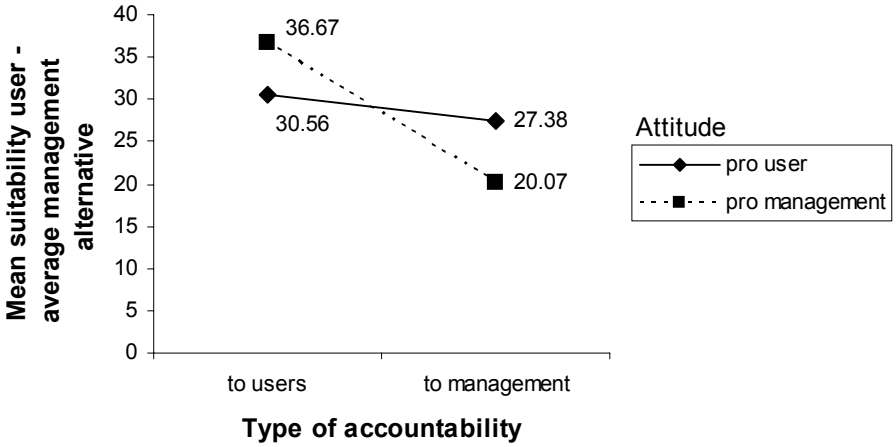
Because the judgmental difference between the user alternative and the average management alternative can be expected to be large if the individual has user preferences and/or has chosen a user alternative, it had been predicted to be larger for participants with a pro-user attitude than with a pro-management attitude, and larger for participants who had been made accountable to the users of the program than for participants who had been made accountable to the management of the company. Also, the difference had been expected to be particularly small for participants with a pro-management attitude who had been made accountable to the management, and particularly large for participants with a pro-user attitude who had been made accountable to the users of the program. Participants experiencing a value conflict between their own preferences and those of the audience they had been made accountable to, however, given that this was an overt measure of differentiation, were expected to signal some adjustment to audience preferences and show a medium large difference. This adjustment was also expected to be greater when participants had been made accountable before the first than before the final decision, resulting in interactions between type of accountability and attitude, time of accountability and attitude, and time of accountability and type of accountability.

In order to test these predictions, 2 (time of accountability: before the first decision vs. before the final decision) by 2 (type of accountability: to users vs. to the management) by 2 (attitude: pro-user vs. pro-management) factorial ANOVAs were conducted on the change in the observed judgmental difference between the user and average management alternative from the first to the final decision as well as the difference obtained after the final decision. The control group was not included in the analysis. The analysis on the change in the judgmental difference between the user and the average management alternative from the first to the final decision did not yield any significant effects. As expected, participants who had been made accountable to the users of the program showed a weak tendency for a larger increase than participants who had been made accountable to the management (M = 9.11 vs. M = 3.72, $F(1,56) = 2.43$, $p = .125$). Also, participants who had been made accountable before the first decision showed a weak tendency for a larger increase than participants who had been made accountable before the final decision (M = 9.78 vs. M = 3.05, $F(1,56) = 2.00$, $p = .163$).

The analysis performed on the judgmental difference between the user and average management alternative observed after the final decision revealed a marginally significant main effect of attitude ($F(1,56) = 2.80$, $p = .100$), a marginally significant interaction between time of accountability and type of accountability ($F(1,56) = 3.66$, $p = .061$) and a weak tendency for an interaction between time of accountability and attitude ($F(1,56) = 2.44$, $p = .124$). As expected, participants with a pro-user attitude tended to show a larger difference than participants with a pro-management attitude (M = 31.80 vs. M = 22.10). Simple main effect analyses also suggested that participants who had been made accountable to the users of the program tended to show a larger difference when they had been made accountable before the first compared to the final decision (M = 33.23 vs. M = 20.21, $F(1,56) = 3.05$, $p = .086$). The opposite tended to be true when participants were accountable to the management (M = 23.27 vs. M = 31.10), but not significantly so. Finally, for participants who had been made accountable before the final decision the judgmental difference between the user alternative and the average management alternative was significantly larger when they had a pro-user attitude than when they had a pro-management attitude (M = 34.44 vs. M = 16.88, $F(1,56) = 5.28$, $p = .025$). This was not the case when participants had been made accountable before the first decision (M = 29.17 vs. M = 27.33). Figure 7.19 shows the mean judgmental difference between the user and average management alternative in the various conditions.

This pattern of results is another indicator of the tendency of participants not to adjust to the preferences of the audience they had been made accountable to when they had already committed themselves to another alternative previously, and when they experienced a conflict between a pro-user attitude and having been made accountable to the management. The fact that the judgmental difference between the user and average management alternative was significantly larger for participants who had been made accountable to the users before the first rather than the final decision, indicates that when participants were made accountable to the users before the final decision, those with a pro-management attitude must have rated alternatives according to their attitude and thereby pulled the obtained mean judgmental difference down. This is supported by the significantly smaller judgmental difference for participants with a pro-management attitude compared to a pro-user attitude, which was only observed when they had been made accountable before the final, but not when they had been made accountable before the first decision. Also, the fact that the obtained judgmental difference between the user and average management alternative was relatively high whenever participants had been made accountable to the management (no matter whether before the first or before the final decision), signals that participants with a pro-user attitude must have been generally reluctant to adjust to the preferences of their audience.

a) Accountability before the first decision



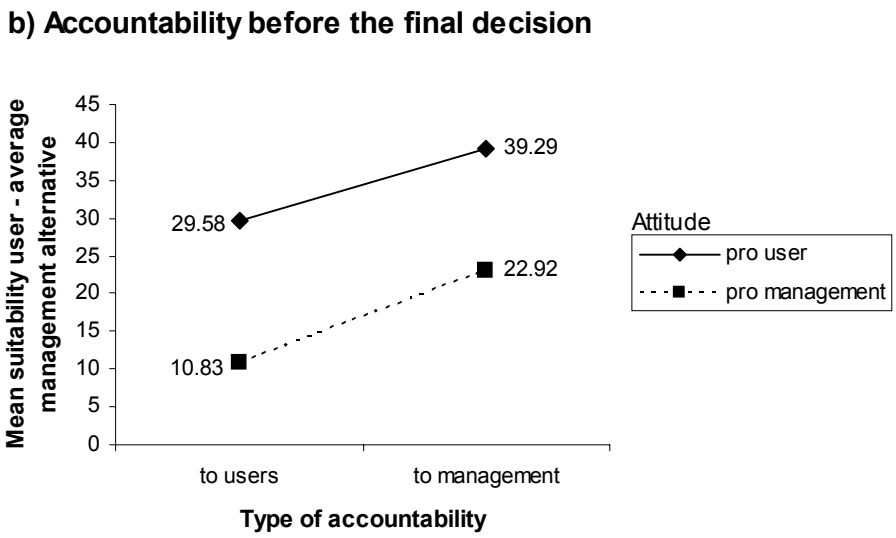


Figure 7.19. Mean difference in overall suitability between the user and average management alternative after the final decision.

3d. MAUT prediction user - average management alternative

The hypotheses concerning the MAUT predicted evaluative differences between the user alternative and the average management alternative had been the following; as for the actually observed judgmental difference between the user and average management alternative, the MAUT differences had been expected to be larger for participants who had been made accountable to the users of the program than for participants who had been made accountable to the management of the company, and larger for participants with a pro-user attitude than with a pro-management attitude. However, in departure from the predictions for the difference in actual ratings, the MAUT predicted evaluative difference between the user and average management alternative had been expected to be particularly *large* for participants with a pro-user attitude who had been made accountable to the management, and particularly *small* for participants with a pro-management attitude who had been made accountable to the users of the program, especially when they had been made accountable before the final decision, resulting in interactions between type of accountability and attitude time of accountability and attitude, and time of accountability and type of accountability.

Again, three-way ANOVAs including the factors of time of accountability, type of accountability and attitude, performed on the data for the experimental groups only, were carried out to test these hypotheses. The analysis of the *change* in the MAUT predicted evaluative difference between the user and average management alternative, based on the attributes from the *first* information set, did not yield any significant effects. There was only a weak tendency for participants with a pro-management attitude to show a larger increase in the MAUT predicted evaluative difference between the user and average management alternative than for participants with a pro-user attitude (M = 1.92 vs. M = -3.89, $F(1,55) = 2.06$, $p = .157$). This suggests that, overall, their preferences had changed towards user preferences.

The ANOVA performed on the change in the MAUT predicted evaluative difference between the user and average management alternative based on the attributes from the *second* information set yielded the expected main effect of type of accountability ($F(1,54) = 6.62$, $p = .013$) and a marginally significant interaction between type of accountability and attitude ($F(1,54) = 2.81$, $p = .100$). Participants who had been made accountable to the management of the company showed a larger increase in the evaluative difference between the user and average management alternative than participants who had been made accountable to the users of the program (M = 13.27 vs. M = 0.82). Simple main effects analyses indicated that this was true only when participants had a pro-user attitude (M = 17.42 vs. M = -2.92, $F(1,54) = 8.97$, $p = .004$). When participants had a pro-management attitude, there was no significant difference (M = 5.06 vs. M = 9.65). This is consistent with the hypothesis that participants with a pro-user attitude who had been made accountable to the management did not adjust to the preferences of their audience, but felt a need to bolster their own preference.

The MAUT predictions based on the combined attributes from *both* information sets showed the same trends that had been observed for the separate attribute sets. The ANOVA carried out on the change in the MAUT predicted evaluative difference between the user and average management alternative from the first to the final decision again yielded a significant main effect of type of accountability ($F(1,53) = 5.34$, $p = .025$) and a weak tendency for an interaction between type of accountability and attitude ($F(1,53) = 2.21$, $p = .143$). As had been observed for the MAUT predictions based on the second attribute set, participants who had been made accountable to the users of the program showed a smaller increase in the evaluative

difference between the user and average management alternative (a decrease, in fact) than participants who had been made accountable to the management of the company (M = -5.90 vs. M = 3.22). Simple main effects analyses, however, indicated that this was true only when participants had a pro-user attitude (M = -8.78 vs. M = 5.70, $F(1,54) = 6.96$, $p = .011$). When participants had a pro-management attitude, the same trend was evident, but not significant (M = -2.48 vs. M = 0.66).

The MAUT predicted differences between the user and average management alternative obtained *after the final decision* showed the following effects. The analysis performed on the predictions based on the attributes from the *first* information set yielded a significant main effect of attitude ($F(1,55) = 7.00$, $p = .011$) and the expected interaction between time of accountability and type of accountability ($F(1,55) = 5.82$, $p = .019$). The mean MAUT predicted evaluative differences between the user alternative and the average management alternative obtained after the final decision are displayed in Figure 7.20. As expected, for participants with a pro-user attitude, the obtained difference was significantly larger than for participants with a pro-management attitude (M = 23.60 vs. M = 11.91). Simple main effects analyses also indicated that when participants had been made accountable before the final decision, the obtained difference was smaller, if they had been made accountable to the users of the program than to the management of the company (M = 14.21 vs. M = 30.81, $F(1,55) = 6.67$, $p = .012$), whereas when participants had been made accountable before the first decision, the observed values were reversed (M = 26.35 vs. M = 18.99), albeit not significantly so. Also, participants who had been made accountable to the management of the company tended to show a smaller MAUT predicted evaluative difference between the user and average management alternative, if they had been made accountable before the first, compared to the final decision (M = 18.99 vs. M = 30.81, $F(1,55) = 3.56$, $p = .065$). For participants who had been made accountable to the users of the program, this tendency was reversed, but not significant (M = 26.35 vs. M = 14.21). Although the triple interaction between time of accountability, type of accountability, and attitude did not reach standard significance levels ($F(1,55) = 1.37$, $p = .247$), simple main effects analyses indicated that the observed interaction between time of accountability and type of accountability was only significant for participants with a pro-management attitude ($F(1,55) = 6.56$, $p = .013$).

This pattern of results suggests that participants who had been made accountable before the final decision did not adjust their unidimensional value and weight judgements to the preferences of the group they had been made accountable to but gave judgements according to their attitude, especially when they experienced a conflict between the two. Accordingly, when participants had been made accountable at the final decision, the obtained evaluative difference between the user and average management alternative was *smaller* when participants had been made accountable to the users and *larger* when they had been made accountable to the management, because for participants with a pro-management attitude who had been made accountable to the users the obtained difference was particularly small and for participants with a pro-user attitude who had been made accountable to the management the obtained difference was particularly large.

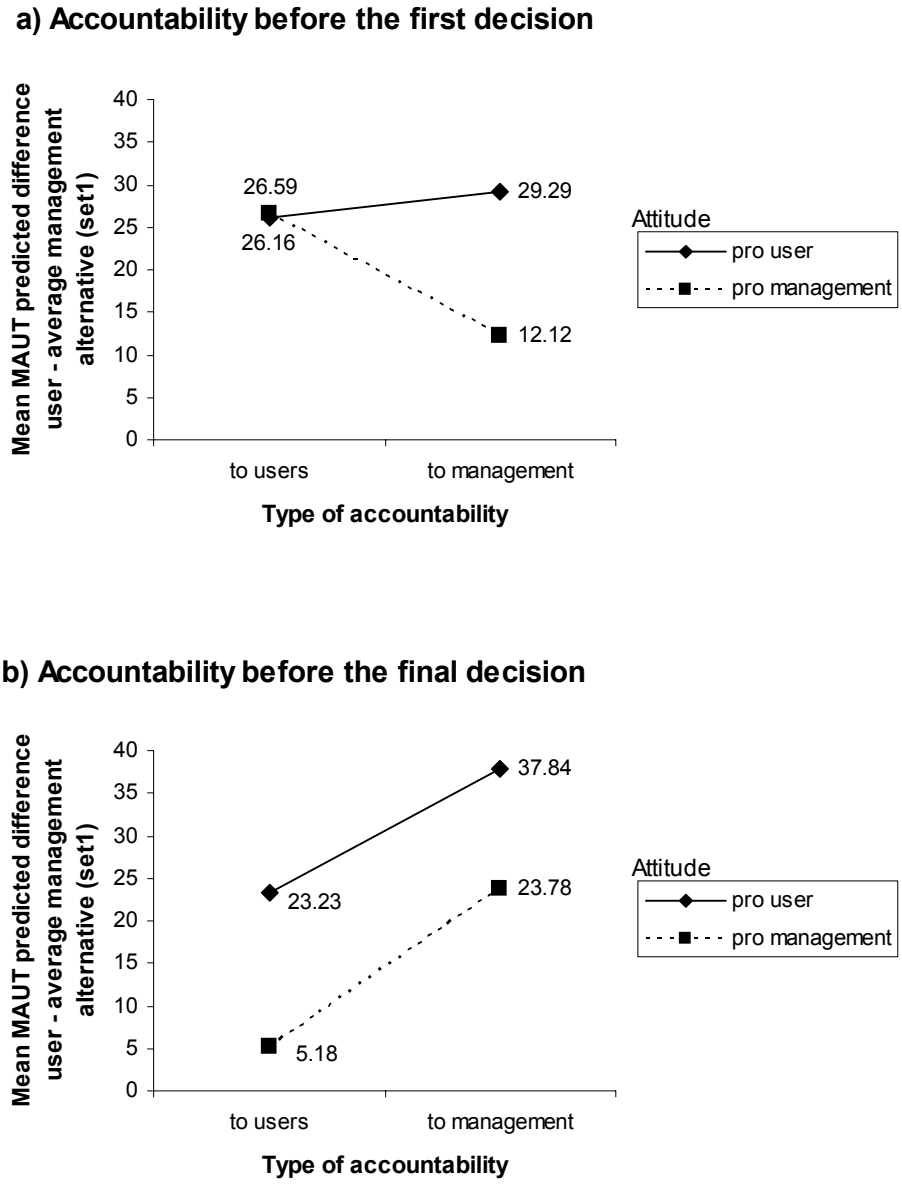


Figure 7.20. Mean MAUT predicted evaluative difference between the user and average management alternative after the final decision (based on the attributes from information set 1).

A similar pattern of results was obtained in the analysis of the MAUT predicted evaluative difference between the user and average management alternative based on the attributes from the *second* information set, except that some of the other main effects were also significant (see Figure 7.21). There was a significant main effect of time of accountability ($F(1,54) = 4.30$, $p = .043$), a marginally significant main effect of type of accountability ($F(1,54)$

= 3.78, p = .057) and a highly significant main effect of attitude (F(1,54) = 14.21, p = .000). The obtained evaluative difference was significantly larger for participants who had been made accountable before the first decision than for participants who had been made accountable before the final decision (M = 40.25 vs. M = 30.28) and for participants with a pro-user attitude compared to a pro-management attitude (M = 45.16 vs. M = 25.37). It also tended to be larger for participants who had been made accountable to the management of the company, compared to the users of the program (M = 40.20 vs. M = 30.64). The interaction between type of accountability and attitude was marginally significant (F(1,54) = 3.04, p = .087). Simple main effects analyses showed that when participants were accountable to the management, a larger evaluative difference was obtained if they had a pro-user attitude than if they had a pro-management attitude (M = 55.90 vs. M = 26.47, F(1,54) = 14.62, p = .000). When participants were accountable to the users of the program, the obtained differences showed the same trend, but were not significantly different from each other (M = 36.32 vs. M = 24.20). Also, for participants with a pro-user attitude the obtained difference between the MAUT predictions for the user and the average management alternative was significantly larger when they had been made accountable to the management of the company than when they had been made accountable to the users of the program (M = 55.90 vs. M = 36.32, F(1,54) = 6.76, p = .012). For participants with a pro-management attitude the obtained difference was not significant (M = 26.47 vs. M = 24.20).

Finally, the analysis performed on the MAUT predicted difference between the user and average management alternative after the final decision, based on the attributes of *both* information set 1 and 2, yielded a marginally significant main effect of type of accountability (F(1,53) = 2.80, p = .100) and a highly significant main effect of attitude (F(1,53) = 14.33, p = .000). Participants who had been made accountable to the management of the company tended to display a larger MAUT predicted evaluative difference between the user and average management alternative than participants who had been made accountable to the users of the program (M = 30.62 vs. M = 24.44). Also, the obtained difference was larger for participants with a pro-user attitude than for participants with a pro-management attitude (M = 35.75 vs. M = 19.28).

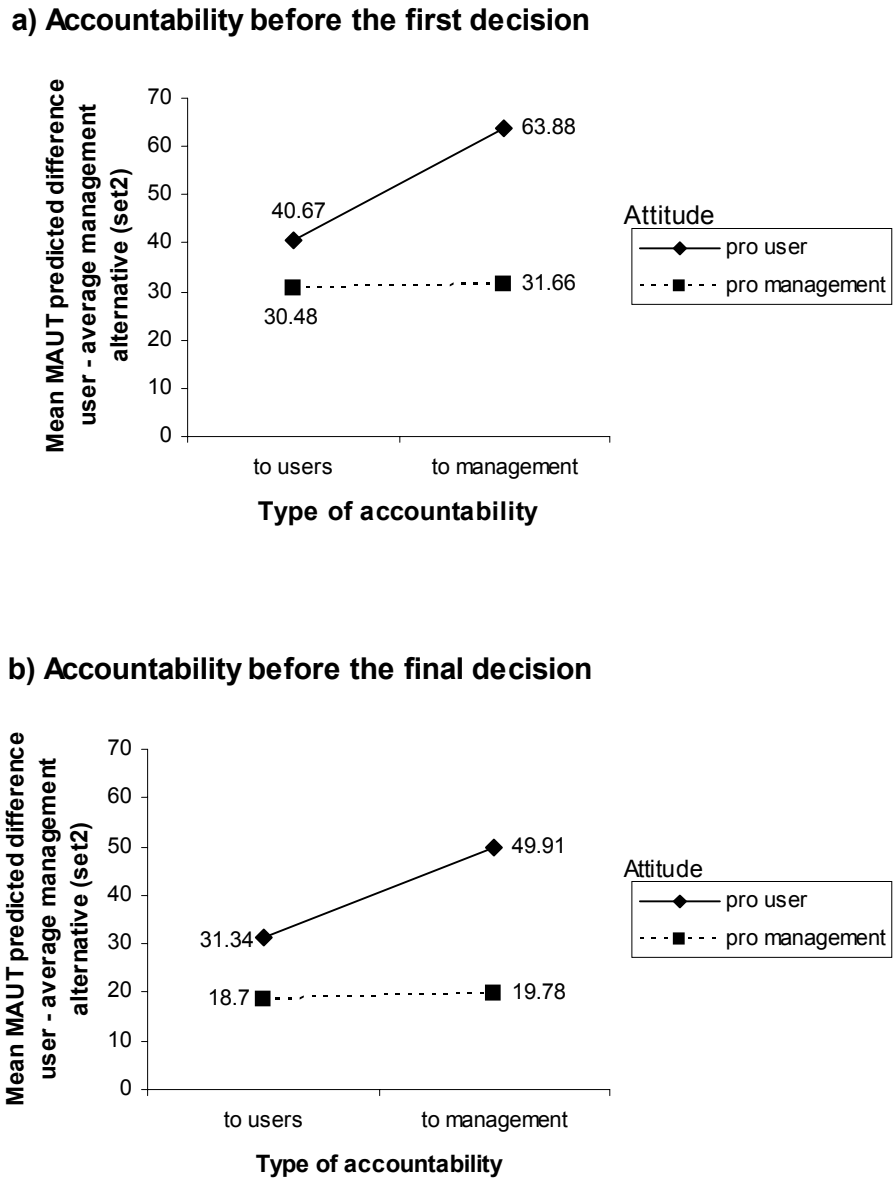


Figure 7.21. Mean MAUT predicted evaluative difference between the user and average management alternative after the final decision (based on the attributes from information set 2).

None of the interactions reached standard significance levels, but a number of significant simple main effects were obtained. When participants had been made accountable before the final decision, those who had been made accountable to the management of the company showed

a larger difference between their MAUT prediction for the user alternative and the average management alternative than those who had been made accountable to the users of the program (M = 32.32 vs. M = 18.37, $F(1,53) = 4.52$, $p = .038$). When participants had been made accountable before the first decision, on the other hand, this difference was not significant (M = 28.78 vs. M = 30.51). Also, when participants had been made accountable to the users of the program, the obtained difference in MAUT predictions between the user alternative and the average management alternative tended to be larger when they had been made accountable before the first compared to the final decision (M = 30.51 vs. M = 18.37, $F(1,53) = 3.63$, $p = .062$) and when they had a pro-user attitude compared to a pro-management attitude (M = 30.32 vs. M = 17.77, $F(1,53) = 3.72$, $p = .059$). For participants who had been made accountable to the management, time of accountability did not significantly affect the obtained difference in MAUT predictions between the user alternative and the average management alternative (accountability before first decision: M = 28.78, accountability before final decision: M = 32.32). There was, however, a highly significant difference between participants with a pro-user attitude and a pro-management attitude ($F(1,53) = 11.42$, $p = .001$); participants with a pro-user attitude who had been made accountable to the management showed a significantly larger difference in MAUT predictions between the user and the average management alternative than participants with a pro-management attitude (M = 42.84 vs. M = 20.69). Finally, for participants with a pro-user attitude, this difference tended to be larger when they had been made accountable to the management of the company compared to the users of the program (M = 42.84 vs. M = 30.32, $F(1,53) = 3.86$, $p = .055$), whereas there was no significant difference for participants with a pro-management attitude (M = 20.69 vs. M = 17.77).

Summary of the results for information evaluation and integration at the final decision

Overall, the results relating to information evaluation and integration processes at the final decision largely confirmed the predictions that had been made at the outset, in that accountability after a commitment had already been made seemed to result in reluctance to adjust to the preferences of the group participants were made accountable to if there was a conflict between their own attitude and the preferences of their audience. Results also suggested that

conflict was experienced as particularly strong if participants had a pro-user attitude and were made accountable to the management.

As expected, participants' choices indicated adjustment to the preferences of the group they had been made accountable to when they had already been made accountable before the first, but not when they had been made accountable before the final decision. The choices of participants who had been made accountable only after they had already made a commitment to one of the alternatives were in accordance with their own attitude, that is, showed a preference for the user alternative when participants had a pro-user attitude, and a preference for management alternatives when participants had a pro-management attitude. Participants who had already been made accountable before the first decision, however, chose according to the preferences of their external audience; they chose the user alternative more often when they had been made accountable to the users of the program, and a management alternative more often when they had been made accountable to the management of the company, irrespective of their own attitude. Surprisingly, in their final choice participants showed an overall preference for the user alternative when they had not been made accountable at any point, even if their own attitude was management-oriented. This was because three participants had changed their initial choice from a management alternative to the user alternative, demonstrating the overall bias towards user considerations in the sample. Otherwise, changes of the preferred alternative were relatively rare and did not vary as a function of experimental condition.

Participants' ratings of attribute importance confirmed the effect evident in their choices, but only where the attributes from the first information set were concerned. In this case, when participants had been made accountable before the first decision, the mean sum of user attribute weights was significantly higher in the 'accountability to users' condition than in the 'accountability to the management' condition, whereas there was no significant difference for participants who had been made accountable before the final decision. Indeed, the sum of user attribute weights was also significantly higher when participants had been made accountable to the management immediately prior to the final decision rather than already before the first decision. This effect may be explained by participants' choices. A majority of participants who had been made accountable to the management before the final decision had chosen the user alternative before and kept this choice, which made it necessary for them to stress the importance

of attributes on which their chosen alternative fared particularly well. Surprisingly, the weight judgements for attributes contained in information set 2 did not show the same effect as those for attributes contained in information set 1. The only significant effect observed was a main effect of attitude; as expected, participants with a pro-user attitude assigned higher user attribute weights than participants with a pro-management attitude. The fact that also no significant effects were obtained for the changes in weight judgements from the first to the final decision, signals that regression effects seem to have been stronger than any importance differentiation and consolidation effects as such. Supporting this assumption is the fact that the grand means were all negative (set 1: M = -.008, set 2: M = -.006, set 1+2: M = -.005).

The actually observed and MAUT predicted evaluative differences between the user alternative and average management alternative also showed some evidence for a reluctance to adjust to audience preferences for participants who had been made accountable only before the final rather than already before the first decision when their own attitude was incompatible with the preferences of the audience they had been made accountable to. In addition, it emerged that a conflict between a pro-user attitude and accountability to the management induced a particularly strong need to bolster the own preference. Although these effects were not significant in the change of the actually observed judgmental difference between the user and average management alternative from the first to the final decision, they were observed in the judgmental differences between the user and average management alternative obtained after the final decision and the MAUT predicted differences observed after the final decision. These evaluative differences between the user and average management alternative were larger when participants had been made accountable before the first rather than the final decision in the 'accountability to users' conditions and larger when participants had been made accountable before the final rather than the first decision in the 'accountability to the management' conditions. This is consistent with a failure to adjust to audience preferences and bolstering of own preferences when being made accountable before the final decision. Also, significantly smaller evaluative differences between the user and average management alternative were observed for participants with a pro-management attitude when they had been made accountable to the management rather than the users, whereas for participants with a pro-user attitude there was no significant difference. This signals that participants with a pro-user attitude must have been reluctant to adjust to the

preferences of a management audience, whereas participants with a pro-management attitude were happy to adjust to the preferences of a user audience.

The obtained changes in the MAUT predicted differences between the user and average management alternative from the first to the final decision only revealed significant effects of a value conflict between a pro-user attitude and accountability to the management when predictions were based on the attributes contained in the second information set and both information sets combined. This finding was somewhat unexpected, as participants' attribute weights as an important sub-component of these predictions had only shown significant effects for the attributes from the first information set. The fact that time of accountability did not significantly interact with other factors where the changes in the MAUT predicted differences between the user and average management alternative from the first to the final decision were concerned suggests that the measure of change is not sensitive enough to uncover such effects.

The evaluative differences between the chosen and average non-chosen alternative, both those actually observed and those predicted by an additive linear MAUT model, also suggested that a conflict between participants' own preferences and those of the group they had been made accountable to resulted in an enlarged difference, compared to conditions that did not experience conflict, particularly when they had been made accountable before the final rather than the first decision, and particularly when the conflict was between a pro-user attitude and having been made accountable to the management. The nature of the triple interaction between time of accountability, type of accountability and attitude observed in the change of the judgmental difference between the chosen and average non-chosen alternative from the first to the final decision indicated that a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in an increased consolidation when participants had been made accountable before the final decision, but in a decreased consolidation when they had already been made accountable before the first decision. This supported the prediction that participants who had been made accountable before the first decision would adjust to the preferences of the group they had been made accountable to and therefore would not need to show an increased difference between the chosen and average non-chosen alternative, whereas participants who had only been made accountable before the final

decision would want to defend their own choice and therefore needed to increase the evaluative difference between their chosen and average non-chosen alternative.

In the judgmental difference between the chosen and average non-chosen alternative observed after the final decision, the effect of conflict was also evident. In addition, the fact that when they had been made accountable to the users of the program, participants with a pro-management attitude did not show a larger judgmental difference between the chosen and average non-chosen alternative than participants with a pro-user attitude, but participants with a pro-user attitude did show a larger difference than participants with a pro-management attitude when they had been made accountable to the management of the company, suggests that only a pro-user attitude in combination with being accountable to the management created a strong conflict, the opposite, however, did not. It is important to note that participants' choices cannot explain the observed results, because in both conflict conditions (pro-user attitude/accountable to management and pro-management attitude/accountable to users), the same overall choice distribution was observed; 9 out of 15 participants chose the user alternative.

The evaluative difference between the chosen and average non-chosen alternative, when predicted by a MAUT model rather than actually observed, nevertheless showed similar effects of conflict. These were only significant, however, for the changes of this evaluative difference from the first to the final decision, and when predictions were based on the first set of attributes or both attribute sets combined rather than the second. Again, a conflict between a pro-user attitude and accountability to the management resulted in significantly larger evaluative differences than no conflict, whereas the conflict between a pro-management attitude and accountability to the users, although showing the same tendency when the predictions were based on the first attribute set, did not do significantly so. Although the triple interaction between the experimental factors was not significant, simple main effects analyses suggested that these effects were only significant for participants who had been made accountable before the first decision. This may be seen as an indicator for continued bolstering of the chosen alternative, even when accountability pressures were not salient anymore. The only significant effect observed in the analysis of the predictions based on the attributes from information set 2 was an attitude main effect; participants with a pro-user attitude generally showed a larger MAUT predicted evaluative difference between their chosen and average non-chosen alternative than

participants with a pro-management attitude. This failure to observe any other effects can be explained by the absence of such effects in the weight judgements that participants had provided for the attributes of information set 2.

As expected, a presence of accountability pressures meant that, compared to no accountability, the overt judgmental difference between the chosen and the average non-chosen alternative was decreased. Where the change of this difference from the first to the final decision was concerned, a marginally significant interaction between accountability and attitude suggested that when participants had not been made accountable, those with a pro-management attitude increased the judgmental difference between the chosen and average non-chosen alternative, whereas those with a pro-user attitude decreased it.

Overall, there was little evidence for the hypothesis that because of the covert nature of the MAUT predicted evaluative differences between the chosen and average non-chosen alternative, accountable participants would show larger such differences than non-accountable participants. The only significant effect of accountability vs. no accountability was observed for the change in the MAUT predicted evaluative difference between the chosen and average non-chosen alternative from the first to the final decision, based on the attributes from the second information set. This was, however, opposite to what was expected; for accountable participants, the increase in the evaluative difference between their chosen and average non-chosen alternative was smaller than for participants who had not been made accountable.

The MAUT predicted evaluative difference between the chosen and average non-chosen alternative at the final decision, based on the attributes from the first information set, revealed a tendency for an interaction between accountability and attitude, such that when participants had not been made accountable, those with a pro-user attitude showed a larger difference than those with a pro-management attitude. Although showing the same tendency, this difference was not significant, when participants had been made accountable.

Justifications

The justifications provided by participants were content-analysed. The unit of analysis was an argument, defined as the smallest unit that provided some meaning in the experimental

context. Two independent judges sorted the arguments with respect to the categories listed below. Inter-judge correlations were moderate to high (r = .72. to r = .95, mean: r = .89). Differences in categorisation were resolved by discussion. The results are presented in Table 7.14. The following dependent variables were analysed.

- *Total number of arguments*

  This number included four subtypes of arguments: (1) direct arguments which were directly related to information presented during the experiment, (2) elaborations, that is, arguments in which participants had elaborated on information presented during the experiment, and which could not be directly inferred from this information, (3) false reproductions of information, where participants' arguments contained incorrect information, and (4) arguments that mentioned evaluation criteria more generally, without referring to specific alternatives.
- *Complexity of argumentation*

  The complexity of argumentation was measured by the ratio of the number of two-sided arguments to the sum of one-sided and two-sided arguments. Two-sided arguments were defined as arguments which mentioned two alternatives, comparing them in some way, whereas one-sided arguments were arguments that only mentioned one alternative. The complexity index ranged between 0 and 1, where 0 indicated that only one-sided arguments had been presented and the complexity of argumentation was low, and 1 indicated that only two-sided arguments had been presented and the complexity of argumentation was high. It was expected that participants who had only been made accountable after they had already made a decision, that is, participants who had been made accountable before the final decision, would show a higher complexity of argumentation than participants who had already been made accountable before the first decision.
- *Type of arguments*

  Arguments were distinguished in terms of whether they were related to user attributes or cost attributes.
- *Consonant vs. dissonant arguments*

  Consonant arguments were arguments which mentioned a positive aspect of the chosen alternative or a negative aspect of a non-chosen alternative, whereas dissonant arguments were arguments which mentioned a positive aspect of a non-chosen alternative or a negative

aspect of the chosen alternative. Percentages were calculated relative to the total number of arguments excluding global evaluation criteria.

1. Number of arguments

It had been expected that the total number of arguments presented in the justifications would be higher for accountable than for non-accountable participants. This hypothesis was tested with a 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVA, which was performed on the total number of arguments. Although accountable participants mentioned more arguments on average than non-accountable participants (M = 7.98 vs. M = 7.06), the main effect of accountability was not significant nor were any of the other effects. ANOVAs performed on the percentages of direct arguments, elaborations, false reproductions and criteria only yielded a significant main effect of attitude on the percentage of direct arguments ($F(1,75) = 5.17$, $p = .026$). Participants with a pro-management attitude presented a higher percentage of direct arguments than participants with a user attitude (M = 64% vs. M = 52%). This attitude main effect was only marginally significant for the percentage of elaborations observed ($F(1,75) = 2.74$, $p = .102$). In this case, participants with a pro-management attitude presented a lower percentage of elaborations than participants with a pro-user attitude (M = 23% vs. M = 29%)

Table 7.14. Mean number of arguments presented (percentages in brackets).

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=8)	Pro Man. (n=8)
1. Number of arguments										
direct	2.75 (43)	5.38 (76)	2.67 (39)	3.71 (63)	3.57 (57)	5.22 (57)	5.00 (65)	3.71 (74)	5.75 (57)	3.88 (51)
elaboration	2.38 (37)	1.38 (16)	2.56 (29)	2.57 (30)	2.43 (29)	2.44 (25)	1.38 (14)	0.71 (10)	3.25 (36)	3.50 (34)
false reproduction	0.38 (06)	0.13 (02)	0.44 (05)	0.29 (07)	0.43 (06)	0.22 (03)	0.13 (01)	0.57 (09)	0.25 (04)	0.00 (00)
criterion	1.25 (14)	0.50 (06)	2.56 (27)	0.00 (00)	0.86 (08)	1.00 (15)	2.00 (19)	0.29 (07)	0.50 (03)	1.25 (14)
total	6.75 (100)	7.38 (100)	8.22 (100)	6.57 (100)	7.29 (100)	8.89 (100)	8.50 (100)	5.29 (100)	9.75 (100)	8.63 (100)
2. Complexity	0.00	0.00	0.16	0.00	0.00	0.16	0.05	0.09	0.06	0.07
3. Type of argument										
user-related	4.25 (63)	4.69 (63)	4.56 (62)	4.57 (75)	4.21 (56)	4.39 (51)	5.75 (71)	3.00 (58)	4.88 (50)	4.69 (58)
cost-related	2.50 (37)	2.69 (37)	3.67 (38)	2.00 (25)	3.07 (44)	4.50 (49)	2.75 (29)	2.29 (42)	4.88 (50)	3.94 (42)

Table 7.14. cont.

	No Accountability		Accountability before the First Decision				Accountability before the Final Decision			
			To Users		To Management		To Users		To Management	
	Pro User (n=8)	Pro Man. (n=8)	Pro User (n=9)	Pro Man. (n=7)	Pro User (n=7)	Pro Man. (n=9)	Pro User (n=8)	Pro Man. (n=7)	Pro User (n=8)	Pro Man. (n=8)
4a. Consonant arguments										
positive aspects of the chosen alt.	5.13 (94)	6.25 (93)	4.33 (76)	6.00 (94)	5.86 (91)	5.44 (69)	4.75 (76)	4.71 (94)	6.25 (87)	6.38 (87)
negative aspects of the non-chosen alt.	0.25 (04)	0.00 (00)	0.33 (06)	0.29 (03)	0.00 (00)	0.89 (11)	1.00 (14)	0.00 (00)	0.75 (03)	0.25 (06)
total	5.38 (98)	6.25 (93)	4.67 (82)	6.29 (97)	5.86 (91)	6.33 (80)	5.75 (90)	4.71 (94)	7.00 (90)	6.63 (93)
4b. Dissonant arguments										
negative aspects of the chosen alt.	0.13 (02)	0.63 (07)	0.67 (12)	0.29 (03)	0.57 (09)	1.00 (13)	0.50 (07)	0.29 (06)	1.25 (07)	0.75 (07)
positive aspects of the non-chosen alt.	0.00 (00)	0.00 (00)	0.33 (06)	0.00 (00)	0.00 (00)	0.56 (07)	0.25 (03)	0.00 (00)	1.00 (03)	0.00 (00)
total	0.13 (02)	0.63 (07)	1.00 (18)	0.29 (03)	0.57 (09)	1.56 (20)	0.75 (10)	0.29 (06)	2.25 (10)	0.75 (07)

Three-factorial ANOVAs tested the effects of time of accountability (before the first vs. before the final decision), type of accountability (to users vs. to the management) and attitude (pro-user vs. pro-management) on the dependent measures for the experimental groups only. The analyses performed on the total number of arguments and the percentage of direct arguments did not yield any effects. For the percentage of elaborations a significant interaction between time of accountability and type of accountability was obtained ($F(1,55) = 4.16$, $p = .046$). Simple main effects analyses indicated that the effect of type of accountability was only significant for participants who had been made accountable before the final decision ($F(1,55) = 6.63$, $p = .013$). When this was the case, participants who had been made accountable to the management showed a significantly higher percentage of elaborations than participants who had been made accountable to the users ($M = 35\%$ vs. $M = 12\%$). For participants who had been made accountable before the first decision, there was no significant difference ($M = 27\%$ vs. $M = 29\%$). Further analyses showed that the type of elaborations that showed this pattern were positive elaborations of the chosen alternative (time of accountability x type of accountability: $F(1,55) = 4.38$, $p = .041$), particularly the user alternative (time of accountability x type of accountability: $F(1,55) = 3.58$, $p = .058$). When participants had been made accountable before the final decision, those who had been made accountable to the management mentioned a significantly higher percentage of positive elaborations of the user alternative than participants who had been made accountable to the users ($M = 15\%$ vs. $M = 4\%$), whereas there was no difference when participants had been made accountable before the first decision ($M = 11\%$ vs. $M = 15\%$, $F(1,55) = 4.15$, $p = .046$). Given that participants who had been made accountable to the management before the final decision often chose a user alternative with bad outcomes on cost attributes, they particularly needed to convince the audience of the benefits of this alternative and seemed to have done so by elaborating on the information provided.

The analysis carried out on the percentage of false reproductions resulted in a marginally significant effect between type of accountability and attitude ($F(1,55) = 3.08$, $p = .085$). Simple main effects analyses suggested that there was a tendency for participants with a pro-management attitude to reproduce a higher percentage of information incorrectly when they had been made accountable to the users than when they had been made accountable to the

management (M = 8% vs. M = 1%, $F(1,55) = 3.08$, $p = .085$). There was no significant difference when participants had a pro-user attitude (M = 3% vs. M = 5%). Further analyses indicated that the falsely reproduced information tended to support the chosen alternative (type of accountability x attitude: $F(1,55) = 3.95$, $p = .052$), in particular the user alternative (type of accountability x attitude: $F(1,55) = 4.88$, $p = .031$). When participants had a pro-management attitude, they tended to mention relatively more information that falsely supported the user alternative when they had been made accountable to the users than when they had been made accountable to the management (M = 3% vs. M = 0%, $F(1,55) = 3.66$, $p = .061$), whereas the opposite, however, non-significant, pattern was observed when participants had a pro-user attitude (M = 1% vs. M = 3%). This seems to suggest that an attitude that was incompatible with the preferences of the audience participants had been made accountable to, especially a pro-management attitude, induced a tendency to report incorrect information that supported the audience preference, possibly in an attempt to make a good impression on the audience.

The interaction between type of accountability and attitude was also significant in the ANOVA testing the effects of the experimental variables on the percentage of arguments that consisted of global evaluation criteria rather than arguments referring to specific alternatives ($F(1,55) = 6.76$, $p = .012$). Participants with a pro-user attitude mentioned a higher percentage of global criteria when they had been made accountable to the users of the program than when they had been made accountable to the management of the company (M = 23% vs. M = 6%, $F(1,55) = 5.08$, $p = .028$). The opposite tended to be true when participants had a pro-management attitude, but the difference failed to reach significance (M = 3% vs. M = 15%). Also, when participants had been made accountable to the users of the program, those with a pro-user attitude mentioned a significantly higher percentage of criteria than those with a pro-management attitude (M = 23% vs. M = 3%, $F(1,55) = 6.16$, $p = .016$), whereas this difference was reversed when participants had been made accountable to the management of the company (M = 6% vs. M = 15%), although not significantly so. Further analyses suggested that this pattern of results was obtained particularly when the importance of user attributes was stressed (type of accountability x attitude: $F(1,55) = 3.11$, $p = .083$). When participants had a pro-user attitude, those who had been made accountable to the users

mentioned a higher percentage of arguments stressing the importance of user criteria than those who had been made accountable to the management (M = 9% vs. M = 2%, $F(1,55)$ = 3.06, p = .086) and than participants with a pro-management attitude (M = 9% vs. M = 2%, $F(1,55)$ = 3.06, p = .086). This indicates an attempt of participants to refer to evaluation criteria that they perceived as important in the situation, could expect their audience to agree with, and on which their chosen alternative tended to excel. This was particularly the case for user criteria and a choice of the user alternative.

2. Complexity of argumentation

The overall complexity of argumentation was very low, indicated by a grand mean of 0.06, where 0 indicates minimal complexity. It had been expected that non-accountable participants would show a lower complexity of argumentation than participants who had been made accountable at some point. The data, analysed with a two-factorial ANOVA including the factors of accountability and attitude, only revealed a weak tendency for this to be the case (M = 0.00 vs. M = 0.08, main effect of accountability: $F(1,75)$ = 2.60, p = .111). It should be noted that among participants who had not been made accountable, not a single one presented any two-sided arguments that compared alternatives, providing strong support for the assumption of a low complexity of argumentation in this condition.

A three-factorial ANOVA testing the effects of time of accountability, type of accountability and attitude on the complexity of argumentation observed in the experimental conditions revealed a marginally significant triple interaction between these factors ($F(1,55)$ = 3.09, p = .085). Simple main effects suggested that a significant interaction between type of accountability and attitude was obtained in the 'accountability before the first decision' conditions only ($F(1,55)$ = 5.34, p = .025). When participants had been made accountable before the first decision, the complexity of argumentation was significantly lower when participants had been made accountable to an audience with preferences incompatible with their own attitude (M = 0.00 for both conditions) than when they had been made accountable to an audience with compatible preferences (M = 0.16 for both pro-user/to users and pro-management/to management). When participants had been made accountable before the final decision, no significant differences between conditions were observed (pro-user/to users: M =

0.05, pro-user/to management: $M = 0.09$, pro-management/to users: $M = 0.06$, pro-management/to management: $M = 0.07$). These relationships are illustrated in Figure 7.22.

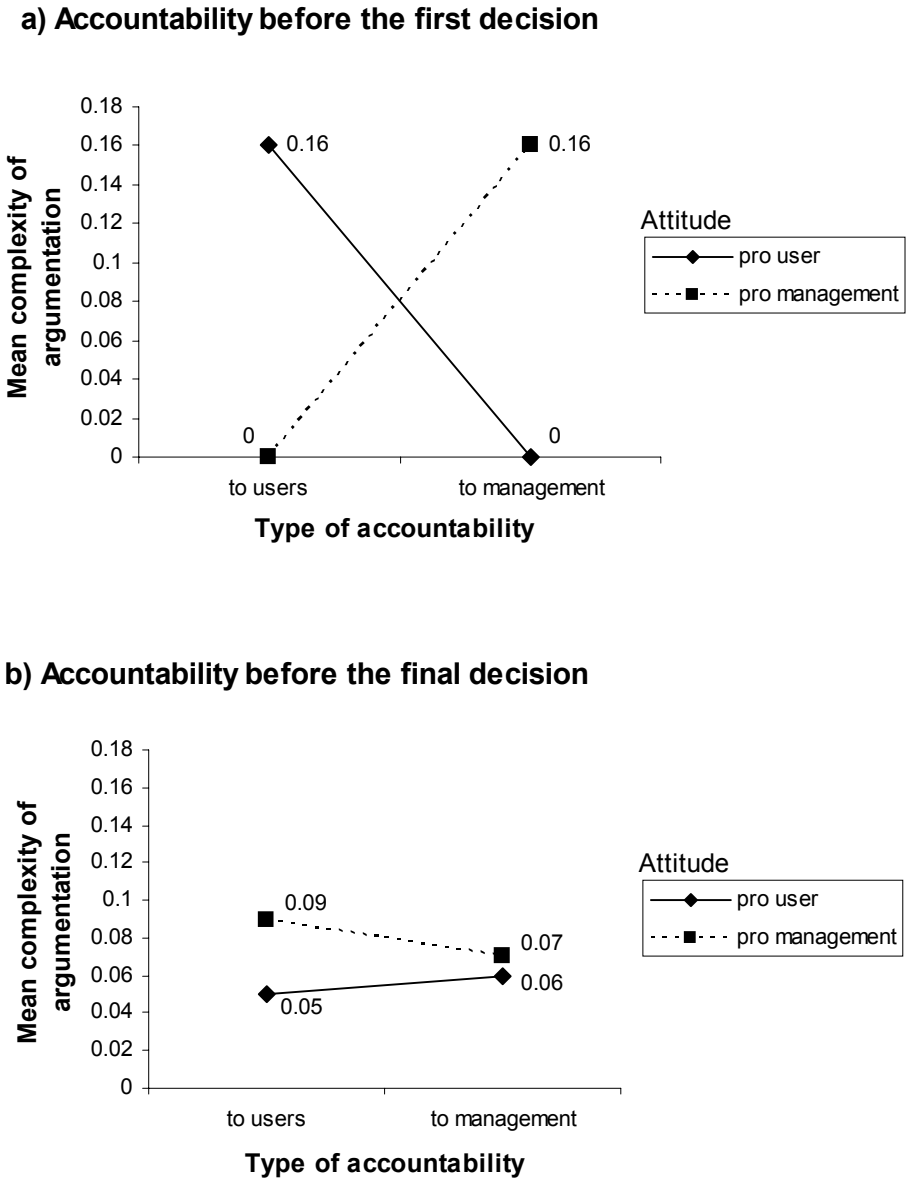


Figure 7.22. Mean complexity of argumentation (1 = maximum complexity).

3. Type of argument

Overall, participants provided a higher percentage of user arguments than cost arguments (M = 62% vs. M = 38%, $t(78) = 5.36$, $p = .000$). An analysis was only performed on the percentage of user arguments presented, as the percentage of cost arguments was the complimentary value to 100. As expected, a three-factorial ANOVA testing the effects of time of accountability, type of accountability and attitude on the percentage of user arguments revealed a marginally significant main effect of type of accountability ($F(1,55) = 3.56$, $p = .064$). In line with predictions, participants who had been made accountable to the users of the program mentioned a higher percentage of user arguments than participants who had been made accountable to the management of the company (M = 66% vs. M = 56%), thereby signalling attention and adjustment to the norms of the group they had been made accountable to.

4a. Consonant arguments

Overall, the percentage of consonant arguments that participants presented was much higher than that of dissonant arguments (M = 92% vs. M = 8%, $t(76) = 27.96$, $p = .000$). Within consonant arguments, participants showed a clear preference for mentioning positive aspects of the chosen alternative compared to mentioning negative aspects of the non-chosen alternatives (M = 88% vs. M = 5%, $t(76) = 24.49$, $p = .000$). Although, as expected, non-accountable participants tended to mention a higher percentage of positive aspects of the chosen alternative (M = 93% vs. M = 86%), negative aspects of the non-chosen alternatives (M = 5% vs. M = 2%) and consonant arguments overall (M = 96% vs. M = 92%) than accountable participants, none of the observed differences were significant.

A three-factorial ANOVA carried out on the percentage of positive aspects of the chosen alternative that was mentioned by participants in the experimental groups yielded a marginally significant interaction between type of accountability and attitude ($F(1,53) = 3.46$, $p = .069$). Simple main effects analyses suggested that the effect of type of accountability only tended to be significant, when participants had a pro-management attitude ($F(1,55) = 2.67$, $p =$

.108). In this case, participants mentioned a higher percentage of positive aspects of the chosen alternative when they had been made accountable to the users of the program than when they had been made accountable to the management of the company (M = 94% vs. M = 81%). When participants had a pro-user attitude, the opposite tended to be true; they mentioned a higher percentage of positive aspects of the chosen alternative when they had been made accountable to the management of the company than when they had been made accountable to the users of the program (M = 89% vs. M = 81%, see Figure 7.23) . This suggests that participants felt a particular need to support their chosen alternative when they had been made accountable to a group with preferences that were different from their own.

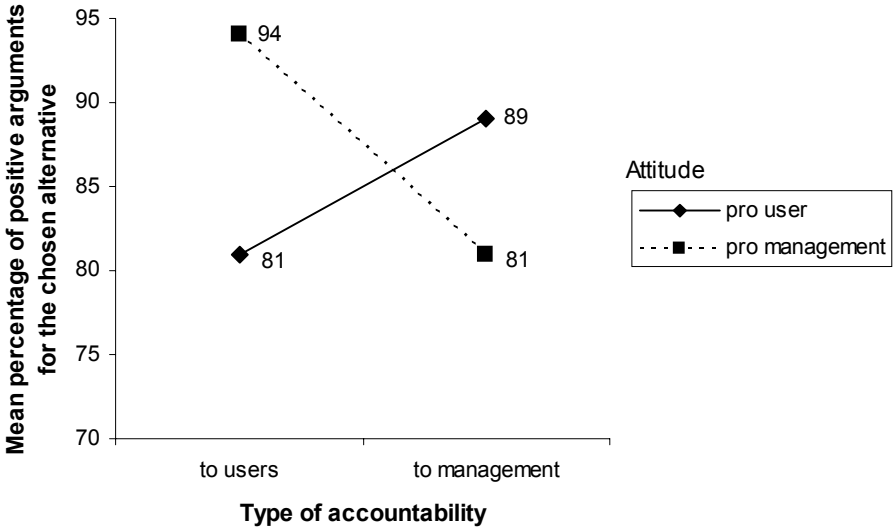


Figure 7.23. Mean percentage of arguments mentioning positive aspects of the chosen alternative, as a function of type of accountability and attitude.

The analysis performed on the percentage of negative aspects of the non-chosen alternatives identified in participants' justifications also revealed a significant interaction between type of accountability and attitude ($F(1,53) = 5.67$, $p = .021$) and a marginally significant interaction between time of accountability and attitude $F(1,55) = 3.03$, $p = .088$). Simple main effects analyses indicated that, whereas when participants had been made accountable to the users of the program, those with a pro-user attitude tended to mention a higher percentage of negative aspects of the non-chosen alternatives than those with pro-

management attitude (M = 9% vs. M = 2%, $F(1,53) = 2.82$, $p = .099$), the opposite tended to be the case when participants had been made accountable to the management of the company (M = 1% vs. M = 8%, $F(1,53) = 2.85$, $p = .097$). Also, for participants with a pro-user attitude, those who had been made accountable to the users of the program tended to mention a higher percentage of negative aspects of the non-chosen alternatives than those who had been made accountable to the management of the company (M = 9% vs. M = 1%, $F(1,53) = 3.06$, $p = .086$). The difference for participants with a pro-management attitude showed the reverse trend (M = 2% vs. M = 8%, see Figure 7.24), but was not significant. Unlike the results obtained for the percentage of positive aspects of the non-chosen alternative, these findings suggest that participants tended to mention relatively fewer negative aspects of the non-chosen alternatives when they expected to justify their choice to an audience with incompatible preferences. This may be explained by the fact that participants showed some tendency to strategically adjust to the preferences of the audience they had been made accountable to. When a conflict between participants' own attitude and the preferences of their audience existed, they may have felt that they could not express their real preference, but would have been nevertheless reluctant to say something negative about it.

The simple main effects analyses performed to explore the marginally significant interaction between time of accountability and attitude did not allow very clear conclusions but suggested that there was a weak tendency for participants with a pro-user attitude to mention a higher percentage of negative aspects of the non-chosen alternatives when they had been made accountable before the final rather than the first decision (M = 8 vs. M = 2, $F(1,53) = 2.56$, $p = .116$). This is consistent with a reluctance of participants with a pro-user attitude in particular to adjust to audience preferences after they had already made a commitment to a different alternative.

Finally, the analysis on the percentage of both types of consonant arguments combined did not reveal any significant effects.

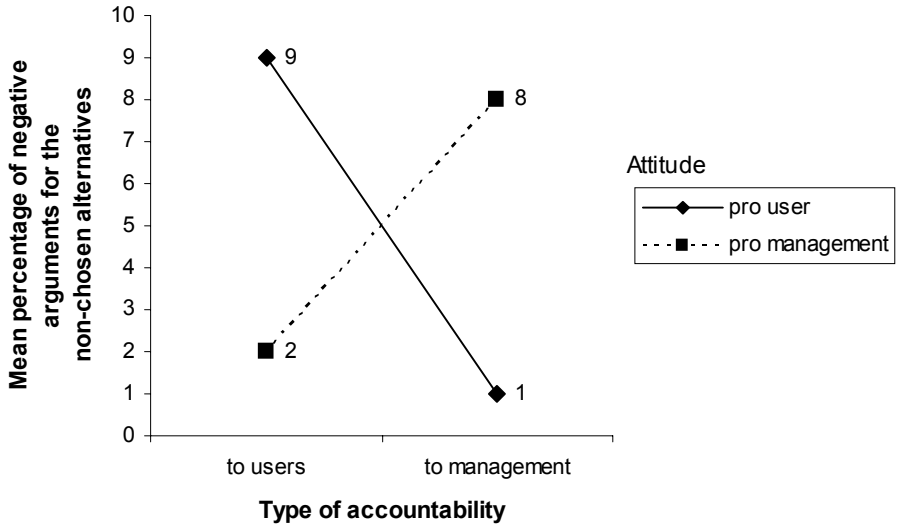


Figure 7.24. Mean percentage of arguments mentioning negative aspects of the non-chosen alternatives, as a function of type of accountability and attitude.

4b. Dissonant arguments

As already mentioned, the occurrence of dissonant arguments was generally rare, but when they were mentioned, participants tended to do so relatively more often in the form of negative aspects of the chosen alternative than positive aspects of the non-chosen alternatives (M = 6% vs. M = 1%, $t(76) = 3.98$, $p = .000$). As expected, non-accountable participants tended to mention a lower percentage of negative aspects of the chosen alternative (M = 4% vs. M = 7%), positive aspects of the non-chosen alternatives (M = 0% vs. M = 2%) and dissonant arguments overall than accountable participants (M = 4% vs. M = 9%), but neither of these differences was significant, as the accountability main effects in the two-factorial ANOVAs including the factors of accountability and attitude that were performed on these measures were not significant. Also, neither of the three-factorial ANOVAs testing the effects of time of accountability, type of accountability and attitude on these measures yielded any significant effects.

Summary of the justification results

Overall, participants' written justifications strongly focused on the advantages of the chosen alternative and exhibited a very low complexity, almost exclusively mentioning one-sided arguments and not comparing alternatives in any way. In line with predictions, non-accountable participants showed this tendency more strongly, and mentioned consonant arguments (positive aspects of the chosen alternative and negative aspects of the non-chosen alternatives) relatively more often and dissonant arguments (negative aspects of the chosen alternative and positive aspects of the non-chosen alternatives) relatively less often than accountable participants, but for none of these dependent measures the observed difference was significant. Non-accountable participants also tended to mention fewer arguments overall than accountable participants, but, again, not significantly so.

The analyses performed on the experimental conditions only, however, revealed some interesting and partly predicted significant effects. The relative frequency of user and cost arguments observed in the written justifications indicated that participants attended to the preferences of the group they had been made accountable to. A marginally significant main effect of type of accountability showed, as expected, that participants who had been made accountable to the users of the program referred to user arguments relatively more often than participants who had been made accountable to the management in their justifications.

When participants had been made accountable before the final decision, the justifications of those participants who had been made accountable to the management included a significantly higher percentage of elaborations, that is, arguments that were not directly based on information provided by the experimenter, than those of participants who had been made accountable to the users. These elaborations particularly supported the user alternative and therefore suggest that, because they chose a user alternative in the majority of cases (10/16), participants who had been made accountable to the management before the final decision needed to bolster their choice with elaborations.

Another effect that emerged was that of participants with a pro-management attitude to reproduce a higher percentage of information incorrectly when they had been made

accountable to the users, compared to when they had been made accountable to the management. The fact that this information tended to support the user alternative may suggest that participants in this condition wanted to make a good impression on their audience when they perceived a conflict between their own attitude and that of their audience. It may also be the case that they really did not remember the advantages of their adopted choice correctly, however, in this case, participants with a pro-user attitude who had been made accountable to the management also should have falsely remembered the advantages of the management alternative. This effect was not observed, however.

Whereas a conflict between a pro-management attitude and having been made accountable to the users of the program resulted in more false reproductions, the absence of this conflict, because participants had a pro-user attitude, seemed to induce a tendency to mention a higher percentage of global evaluation criteria. As expected, these were primarily arguments that supported the importance of user concerns. An equivalent effect was not found for participants with a pro-management attitude who had been made accountable to the management. This may be explained by the fact that for them both user and cost attributes were important, and they tended to choose compromise alternatives that did not have clear advantages on one or the other.

The analyses testing the effects of the experimental variables on the different types of consonant and dissonant arguments only yielded significant effects for consonant arguments. These suggested that a bolstering of the chosen alternative in the form of mentioning a high percentage of positive aspects of the chosen alternative, not surprisingly, particularly occurred when participants perceived a conflict between their own attitude and the preferences of the audience they had been made accountable to. However, where the percentage of negative aspects of the non-chosen alternatives was concerned, a different pattern emerged. This percentage was higher when participants did not experience a value conflict. A possible explanation for these findings is that participants who did perceive a value conflict, showed some tendency to adjust their choice to that they expected to be preferred by the audience they had been made accountable to, especially when they had a pro-management attitude, and especially when they had been made accountable before the first decision. Although they

showed this adjustment, it may be assumed that they would have been reluctant to say something negative about their real preference, resulting in the observed effect.

Finally, the significant triple interaction between the experimental variables obtained in the ANOVA performed on the complexity of the argumentation suggested that the tendency to bolster one's own choice particularly strongly when a value conflict was perceived, was evident especially when participants had been made accountable before the first rather than the final decision. In this case, the complexity of argumentation was significantly lower when participants had been made accountable to an audience with preferences incompatible with their own attitude than when they had been made accountable to an audience with compatible preferences. When participants had been made accountable before the final decision, on the other hand, no significant differences between conditions emerged. This may be regarded as somewhat surprising, as the conflict should have been experienced more strongly if participants had already committed themselves to a different alternative when they were made accountable. On the other hand, accountability before the final decision was more salient to participants, and this finding is therefore in line with the expected increased depth and complexity of information processing induced by accountability.

Post-experimental questionnaire

1. Difficulty of the decision

It was expected that participants who had been made accountable would find the decisions they had made more difficult than participants who had not been made accountable. A two-factorial ANOVA testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro-management) revealed the expected significant main effect of accountability ($F(1,76) = 4.71$, $p = .033$). Participants who had been made accountable before the first decision indeed found this decision significantly more difficult than non-accountable participants ($M = 3.94$ vs. $M = 3.33$). There was also a significant main effect of attitude ($F(1,76) = 6.86$, $p = .011$). Participants with a pro-management attitude found the decision

more difficult than participants with a pro-user attitude (M = 3.95 vs. M = 3.20). The interaction between accountability and attitude was not significant.

Although accountable participants also rated the final decision as more difficult than non-accountable participants (M = 4.09 vs. M = 3.75), a two-factorial ANOVA testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro- management) on the rated difficulty of the final decision did not yield the expected main effect of accountability nor any other significant effect. A 2 (time of accountability: before the first decision vs. before the final decision) by 2 (type of accountability: to users vs. to the management) by 2 (attitude: pro-user vs. pro-management) ANOVA conducted on the experimental groups only revealed a significant attitude main effect. Participants with a pro-management attitude found the final decision more difficult than participants with a pro-user attitude (M = 4.53 vs. M = 3.66). No other effects were significant.

A repeated-measurements ANOVA with time of rating (after the first vs. after the final decision) as a within-subjects factor and experimental condition and attitude as between-subjects factors revealed a significant 'time of rating' effect ($F(1,70) = 6.41$, $p = .014$). The final decision was rated as more difficult than the first decision (M = 4.02 vs. M = 3.58). The only other significant effect was an attitude main effect ($F(1,70) = 7.12$, $p = .009$). Again, participants with a pro-management attitude generally found the decisions more difficult than participants with a pro-user attitude (M = 4.15 vs. M = 3.45).

2. Confidence in the decision

It had been expected that being made accountable for their decision would reduce participants' confidence in their decision as, for example, found by Tetlock and Kim (1987) and McKenna and Myers (1997). However, two-factorial ANOVAs testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro-management) on reported confidence showed neither any significant effects for the first decision nor for the final decision. There were also no significant effects when the data for participants in the experimental conditions only were analysed.

With respect to a change of confidence from the first to the final decision, it had been expected that participants would report a higher confidence after the final decision than after the first decision, replicating research that has shown that people become increasingly confident in their decision the more information they can base it on (e.g., Oskamp, 1965). A repeated-measurements ANOVA with time of rating (after the first vs. after the final decision) as a within-subjects factor and experimental condition and attitude as between-subjects factors revealed a marginally significant 'time of rating' effect ($F(1,70) = 2.70$, $p = .105$) which, however, if anything, suggested the opposite. Confidence ratings after the first decision were higher than confidence ratings after the final decision ($M = 5.24$ vs. $M = 5.01$).

3. Need to make a good decision

Responses to this question had been expected to vary as a function of whether participants were made accountable or not and the time at which they were made accountable. It was assumed that accountable participants would report a greater need to make a good decision, both at the first and at the final decision, than non-accountable participants, and that participants who had been made accountable before the final decision would report a greater need to make a good final decision than participants who had already been made accountable before the first decision. These assumptions were not supported, however. At the first decision, accountable participants reported a slightly higher need to make a good decision than non-accountable participants ($M = 6.09$ vs. $M = 5.94$), but the main effect of accountability in the 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVA was not significant, nor were any other effects. The same was true for need to make a good decision reported at the final decision ($M = 6.08$ vs. $M = 5.88$). Again, there were no other significant effects. Also, a three-factorial ANOVA testing the effects of time of accountability, type of accountability and attitude on the ratings provided by participants in the experimental groups after the final decision did not yield any significant effects.

A repeated-measurements ANOVA with time of rating (after the first vs. after the final decision) as a within-subjects factor and experimental condition and attitude as between-subjects factors, showed that there was also no significant change in ratings from the first to

the final decision (M = 6.00 vs. M = 6.04), as the main effect of time of rating was not significant. Generally, the obtained mean values were quite high (means around 6 on a scale ranging from 1 to 7), suggesting that participants had possibly interpreted the question as one of asking about how seriously they had taken the experiment and had given high ratings in order to make a good impression on the experimenter.

4. Expectation of having to justify the final decision

The assumption that accountable participants would expect more strongly to be asked to justify their final decision than participants who had not been made accountable was not supported, since no significant main effect of accountability was found in the ANOVA employed to test this hypothesis. On the contrary, accountable participants reported a slightly lower expectation than non-accountable participants, although this difference was not significant (M = 2.91 vs. M = 3.00). No other significant effects were found. A three-factorial ANOVA testing the effects of time of accountability, type of accountability, and attitude on the ratings of participants in the experimental groups only also did not show any significant effects. The mean values around 3 indicate that participants' expectations of having to justify their decision were only moderate. This throws some doubt on whether participants in the accountability conditions took the manipulation seriously, but could have also been caused by the format of the justification, which was different from what accountable participants had been told before.

5. Need to be able to justify the first and the final decision

It had been assumed that participants would report a particularly strong need to justify their decision immediately after being made accountable, resulting in a significant difference in ratings of this need between accountable and non-accountable participants for the first decision and between participants who had been made accountable before the first decision compared to participants who had only been made accountable before the final decision for the final decision. These hypotheses were not confirmed.

There was no significant difference in the reported need to be able to justify the *first* decision for accountable and non-accountable participants, as the main effect of accountability in the ANOVA carried out to test this hypothesis was not significant. The main effect of attitude and the interaction between accountability and attitude were also not significant. A three-factorial ANOVA performed on the experimental groups only, however, revealed a significant three-way interaction between time of accountability, type of accountability, and attitude ($F(1,56) = 4.51$, $p = .038$). A simple main effects analysis indicated that the interaction between time of accountability and attitude was only significant in conditions where participants had been asked to justify their decision to the management rather than the users ($F(1,56) = 4.07$, $p = .049$). Participants who had been made accountable to the management of the company before the first decision reported a higher need to justify this decision when they had a pro-management attitude compared to a pro-user attitude ($M = 5.33$ vs. $M = 4.29$). For participants who were only made accountable to the management at a later time, this pattern was reversed ($M = 4.63$ vs. $M = 5.38$). This relationship is illustrated in Figure 7.25.

Assuming that participants were willing and able to report their perceived need to justify their first decision accurately, this suggests that accountability to the management for participants with a pro-management attitude created a particular pressure to perform well. Participants who had not been made accountable yet at the time, on the other hand, seem to have felt more under pressure when their own attitude was user-oriented than when it was management-oriented.

Accountable and non-accountable participants did also not differ in their reported need to justify the *final* decision, since the accountability main effect in the 2 (accountability: yes vs. no) by 2 (attitude: pro-user vs. pro-management) ANOVA conducted on this measure did not show a significant accountability main effect, nor any other significant effects. This was also true for the three-factorial ANOVA which tested the effects of time of accountability, type of accountability and attitude on the ratings of the experimental groups only.

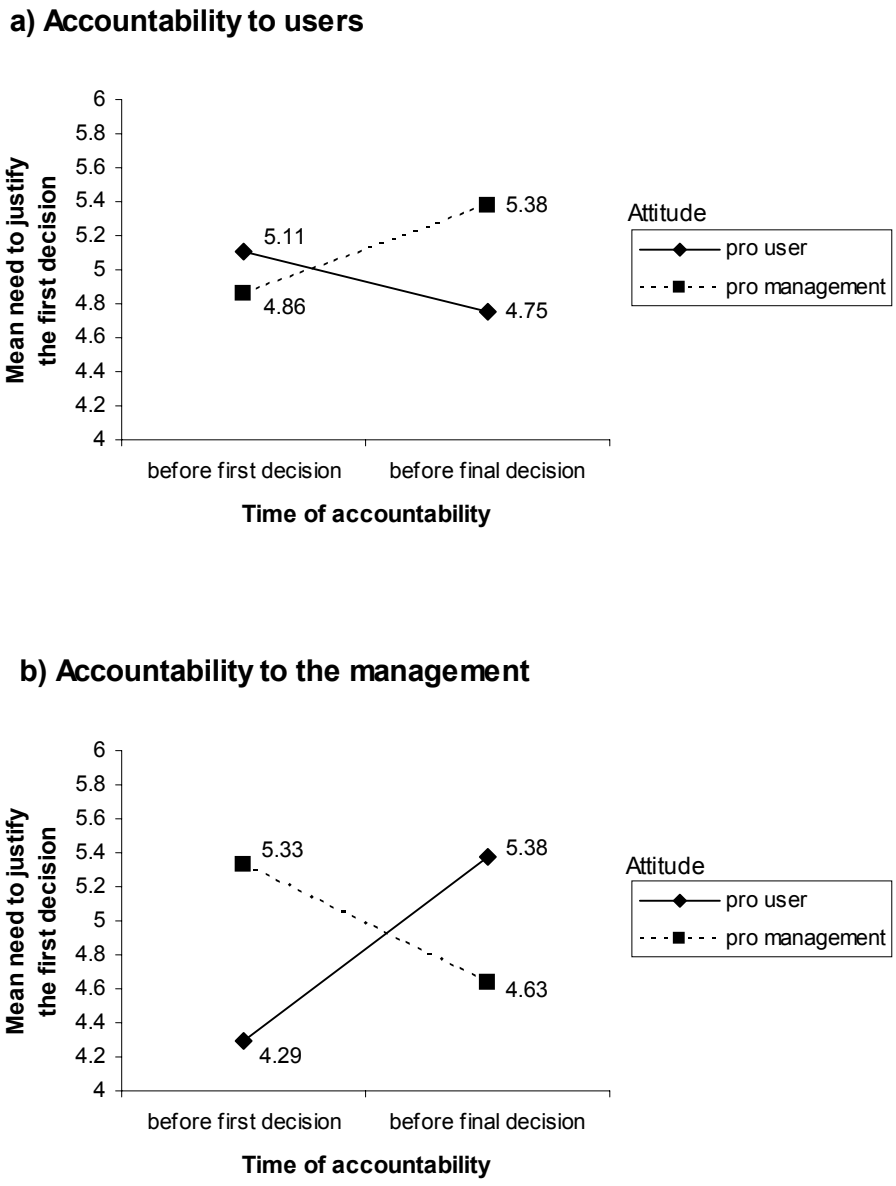


Figure 7.25. Mean need to be able to justify the first decision (1=not at all, 7=extremely).

A repeated-measurements ANOVA with type of decision (first vs. final decision) as a within-subjects factor and experimental condition and attitude as between-subjects factors revealed a significant 'type of decision' effect ($F(1,70) = 19.43$, $p = .000$). Generally, participants reported a stronger need to justify the final decision than the first decision ($M = 5.66$ vs. $M = 5.05$). No other effects were significant.

6. Usefulness of additional information

It was expected that, generally, participants would find the additional information useful, despite its redundancy, and accountable participants more so than non-accountable participants. The latter was not the case. The two-factorial ANOVA testing the effects of accountability (yes vs. no) and attitude (pro-user vs. pro-management) on rated usefulness did not yield the expected significant main effect of accountability, nor any other significant effect. The information was considered to be highly useful, however, indicated by an overall mean value of 5.40, on a scale ranging from 1 ‘not at all useful’ to 7 ‘very useful’.

A three-factorial ANOVA testing the effects of time of accountability, type of accountability and attitude on the usefulness ratings of participants in the experimental groups revealed a marginally significant interaction between time of accountability and type of accountability ($F(1,56) = 3.78$, $p = .057$). A simple main effects analysis showed that type of accountability only significantly affected participants who had been made accountable before the first decision ($F(1,56) = 4.03$, $p = .049$). Participants who had been made accountable before the first decision found the additional information more useful when they had been made accountable to the users of the program than when they had been made accountable to the management of the company ($M = 5.94$ vs. $M = 5.06$), whereas the opposite tended to be the case when participants had been made accountable before the final decision ($M = 5.19$ vs. $M = 5.50$), although not significantly so.

Summary of the results for the post-experimental questionnaire

Participants' responses to the post-experimental questionnaire did not consistently show the expected effects, particularly not where the reported need to be able to justify the first and final decision was concerned. It had been expected that accountable participants would report a greater need to justify their first and final decision, respectively, and that the reported need to be able to justify a decision would be particularly strong if participants had been made accountable immediately prior to that decision. None of these hypotheses were confirmed. Instead, it was found that participants who had been made accountable to the management of the company before the first decision reported a stronger need to justify this

decision when they had a pro-management rather than a pro-user attitude, whereas for participants who had not been made accountable to the management yet, a stronger reported need to justify the first decision was observed when they had a pro-user rather than a pro-management attitude. This may be seen as an indicator for the fact that for participants with a pro-management attitude, accountability to the management at the first decision created a particular pressure to perform well. For participants with a pro-user attitude who had been made accountable to the users, this effect was not observed. A repeated-measurements ANOVA found that, generally, participants indicated a stronger need to justify the final than the first decision, which is not surprising, given that the final decision was announced as such and therefore must have appeared as particularly important.

Accountability also did not significantly affect the reported need to make a good first or final decision, although the observed means were higher for accountable than non-accountable participants and therefore in the predicted direction. There was no difference between the first and final decision with respect to the reported need to make a good decision; the generally high values that were observed suggest that participants may have interpreted these questions as testing their motivation to comply with experimental instructions and responded in the affirmative, in order to make a good impression.

Participants generally only reported a moderate expectation of having to justify the final decision, which did not vary significantly as a function of accountability. This may mean that the accountability manipulation was not convincing enough, but could also have been caused by the format of the justification; originally, accountable participants had expected to justify their choice to other people in a face-to-face discussion, but then found themselves asked to provide a written justification.

As predicted, accountable participants found the first decision more difficult than non-accountable participants. The same trend was observed for the final decision, but failed to reach statistical significance. Generally, the final decision was perceived as more difficult than the first, which is consistent with the result reported earlier that participants perceived a stronger need to be able to justify it than they did for the first decision. The observed ratings of difficulty also revealed significant main effects of attitude. Participants with a pro-

management attitude perceived both the first and the final decision as more difficult than participants with a pro-user attitude. This supports the conclusion suggested by the information search results in particular, namely, that participants with a pro-management attitude took both user and cost attributes into account when making their decision. This put a higher strain on their information processing capacities compared to participants with a pro-user attitude, and therefore made decisions indeed more difficult for them.

Unexpectedly, there was neither a significant reduction of confidence in the first nor in the final decision as a result of accountability. Confidence was also not increased after the final decision, as had been predicted originally. If anything, an opposite trend could be observed; participants tended to be less confident after the final than the first decision.

Finally, the redundant information that was presented to participants in the second information set was perceived as highly useful overall. The perceived usefulness did not vary as a function of accountability as such, but analyses revealed a significant interaction between time of accountability and type of accountability in the experimental groups. Participants who had been made accountable before the first decision rated the additional information as more useful when they had been made accountable to the users of the program than when they had been made accountable to the management of the company, whereas a reverse (but not significant) trend was observed for participants who had only been made accountable before the final decision. Possibly, participants who been made accountable to the management before the first decision and had felt under pressure to justify their decision, especially if they had not adjusted to the preferences of their audience, may have particularly hoped to find more information that supported their choice. This was, however, not possible, because of the redundancy of this information with the information provided in the first set, and these participants may have therefore been particularly disappointed.

Discussion and conclusions

The present study had been designed to test the norm-enforcement aspect of accountability, as predicted by the PMA presented in Chapter 5. According to the PMA, external accountability to others who are able to evaluate the individual, because his or her performance is identifiable, and whose standards of evaluation have been made explicit to or can be easily guessed by the individual, will make directional goals salient and bias the decision process towards the alternative that is likely to be preferred by the individual's evaluators. However, this is only predicted to be the case if the decision maker has not already committed him- or herself to a different alternative, that is, accountability is pre- rather than post-decisional. When the individual is made accountable after having made a decision already, he or she is expected to evaluate and integrate information in such a way that the previously chosen alternative comes out best. Participants' own preferences are assumed to be another important moderator of the tendency to either adjust or not adjust to audience preferences. If they match those of the group the individual has been made accountable to, adjustment to audience preferences is assumed to be more likely compared to when this is not the case.

At the same time as biasing the decision process, the responsibility cues inherent in external accountability demands, namely personal causality in bringing about and control over the decision and its consequences, are predicted to make accuracy goals salient. These are expected to enhance depth of information search and complexity of information integration. External accountability to representatives of different interest groups was therefore predicted to make the decision process both more complex and more biased.

These hypotheses were tested in a decision context in which participants had to choose which of four CASE programming tools to buy for a company and were made accountable either to the future users of the program (the programmers of the company) or to the management of the company. This happened either before or after an initial decision was made. Participants in a control group were not made accountable at any time. The accountability manipulation was implicit, in that participants were told that the final decision

about which CASE tool to acquire would be made by a committee, consisting of themselves and two other people, a representative of the users and a representative of the management. Participants were lead to believe that, after having made a decision on their own, they would meet these representatives and have to jointly agree on a decision after a group discussion. Accountability to the users of the program was created by implying that it would be particularly important to convince the user representative, as the users would have to work with the tool in the future, accountability to the management was created by implying that it would be particularly important to convince the management representative, as the management would have to pay for it. In this way the values of the audience participants were made accountable to were made salient, and it was expected that participants who had been made accountable to the users of the program before they made a decision would bias their decision process towards an alternative that maximised user values, whereas participants who had been made accountable to the management of the company would bias their decision process towards an alternative that maximised cost values.

However, the findings suggested that participants who had been made accountable to the management of the company before they made a decision attempted to maximise *both* user and cost values, which meant that the results did not usually exactly match the original predictions, but nevertheless provided evidence for the type of processes predicted by the PMA. Another tendency that affected results was the strong user orientation in the sample. This may not seem surprising, given that the participants were all Computer Science students, who were particularly familiar with the user aspects of computer software. This meant, however, that the adjustment to a user audience for participants who had been classified as having a pro-management attitude was easier than the opposite adjustment.

Information search before the first decision

Because of the two effects mentioned above, an increased depth of information search, assumed to be induced as a result of accountability compared to no accountability, could only be observed when participants had a pro-management attitude and had been made accountable to the management. When this was the case, participants indeed exhibited a tendency to adapt

their information search to audience preferences by searching less information when made accountable to the users of the program, and more information when made accountable to the management of the company. This adjustment tendency was also evident in the type of information that was searched; when participants with a pro-management attitude were made accountable to the users of the program, they tended to search relatively more user information, when they were made accountable to the management of the company, they tended to search relatively less user information than under no accountability. As expected, information search in the 'no accountability' condition followed participants' attitude. Because participants with a pro-management attitude had to search both user and cost attributes in order to identify the alternative that offered the best compromise overall, they searched more information than participants with a pro-user attitude when accountability pressures were absent. Also in line with predictions, participants with a pro-user attitude who had been made accountable to the users of the program, tended to search more information, particularly a higher percentage of different cells, than when they were not accountable. Unexpectedly, however, the additional information that was searched was not primarily user information, as the percentage of searched user information did not differ significantly between the 'no accountability' and 'accountability to users' conditions. Participants with a pro-user attitude who had been made accountable to the management searched particularly little information. This points to the validity of an assumption mentioned earlier, namely that for these participants adjustment to a management audience was very difficult. Indeed, the overall pattern of results observed for this condition strongly suggests that an adjustment to the management audience either did not happen or only happened reluctantly.

The observed concentration of the first information search on the chosen alternative, for example, was another indicator of the fact that the conflict between having a pro-user attitude and being accountable to the management was resolved by defending user preferences, whereas the conflict between having a pro-management attitude and being accountable to the users of the program was resolved by adjusting to user preferences. Participants with a pro-user attitude who had been made accountable to the management showed a higher concentration of search on their chosen alternative than participants with a pro-management attitude. For participants with a pro-management attitude who had been

made accountable to the users of the program, however, the equivalent tendency was not significant.

Contrary to predictions, accountable participants did not only fail to show a significantly lower concentration of their information search on the chosen alternative than non-accountable participants, but even exhibited the opposite, albeit non-significant, tendency. Particularly when concentration of search on the chosen alternative was calculated for the total number of items searched, the observed mean was higher for accountable than for non-accountable participants. The fact that this effect was not observed when the concentration of search measure was based on the different cells searched suggests that accountable participants had a stronger need to make sure that their intended choice was superior when they searched information repeatedly, and as such is consistent with more salient accuracy goals for these participants.

The high variability of search as well as the low Index of Compensatory Processing and attribute-wise search pattern indicated that, overall, information search was non-compensatory. It had been predicted that the extent of non-compensatory information processing would be higher for non-accountable than accountable participants, but, although the observed means mostly showed the expected tendencies, none of the observed differences were significant. A possible explanation for this can be the strong influence of task factors, in particular the relatively high complexity of the task, which may have overshadowed any accountability effects. The only significant effect was a main effect of attitude on the variability of search across attributes for different cells. Participants with a pro-management attitude showed a significantly lower variability of search across attributes for different cells than participants with a pro-user attitude. This implies that they searched alternatives more equally and did not eliminate them from consideration as early as participants with a pro-user attitude did. This may be interpreted as supporting the assumption that participants with a pro-management attitude paid attention to both user and cost attributes. The alternatives on offer had been constructed in such a way that they were similar in terms of their overall value and, therefore, someone who was interested in maximising overall value, as participants with a

pro-management attitude were, would have found it hard to eliminate any of them from consideration early.

Information evaluation and integration at the first decision

The measures of information evaluation and integration at the first decision lent further support to the above conclusions. Also, another interesting regularity emerged, namely that overt measures of adjustment, such as choices, weight judgements and suitability ratings of alternatives indicated adjustment to the preferences of the group participants had been made accountable to, whereas covert measures, such as MAUT predicted evaluative differences between alternatives demonstrated the defensive tendencies of participants' whose own attitude conflicted with that of their audience.

Hence, participants who had been made accountable to the users of the program chose the user alternative more often than a management alternative, whereas the opposite was true for participants who had been made accountable to the management. Also, in line with predictions, attitude only significantly affected participants' choice when they were not accountable; as expected, participants with a pro-user attitude were more likely to choose the user alternative and participants with a pro-management attitude were more likely to choose a management alternative. The fact that the pure management alternative was only chosen once by the latter again supports the conclusion that participants with a pro-management attitude were attracted to alternatives that fared well on both user and cost attributes.

Participants' weight judgements reflected their choices. Participants who had been made accountable to the users of the program assigned relatively higher weights to the user attributes of the first information set than participants who had been made accountable to the management, no matter what their own attitude was. The weight judgements of participants in the 'no accountability' condition, on the other hand, were affected by their attitude and, hence, relatively higher for user attributes when participants were user-oriented than when they were management-oriented. As expected, these effects were only observed for the user attributes contained in the first information set, for which participants knew how outcomes were linked to the different alternatives. This suggests that the weight judgements for user attributes

contained in information set 1 were strategically adjusted, in order to support the chosen alternative. The weight judgements for the user attributes of the second information set, on the other hand, only showed the expected main effect of attitude.

Despite the clear adjustment effects that were observed in participants' weight judgements, the change of these judgements from the first elicitation (which had happened before any experimental manipulation) to the second elicitation after the first decision, although showing the predicted trends, did not reveal any significant effects, suggesting that effects were not strong enough to be able to speak of importance *differentiation* on an individual basis. An inspection of the mean sum of user attribute weights obtained before the first decision suggested that the reason for the failure to obtain significant differentiation effects may have been the fact that user attribute weights were quite high generally, so that regression effects may have cancelled out any differentiation attempts, especially for participants who had been made accountable to the users of the program and therefore would have had to increase their user attribute weights even more.

Measures of overall attractiveness differentiation between alternatives had included the difference between suitability ratings for the user and the average management alternative, in order to assess the effect of having been made accountable to different interest groups, and the difference between suitability ratings for the chosen and average non-chosen alternative, in order to test differentiation to protect the chosen alternative more generally. Given that both measures were overt measures, in that participants openly expressed their preferences in these ratings, it was expected that participants who had been made accountable would *decrease* the evaluative difference between the chosen and average non-chosen alternative and the user and average management alternative, respectively, in order to signal that they were willing to take differing audience view points into account, especially if they experienced a conflict between their own attitude and the preferences of the group they had been made accountable to.

Indeed, the observed evaluative difference between the user and average management alternative was smaller for participants who had been made accountable to the management than for participants who had either been made accountable to the users of the program or participants who had not been made accountable, suggesting some adjustment effect. This

tendency was not statistically significant, however. There was, however, a significant effect of attitude, which, as expected, indicated that the evaluative difference between the user and average management alternative was larger for participants with a pro-user attitude than for participants with a pro-management attitude. A conflict between participants' own attitude and the preferences of the group they had been made accountable to tended to result in a more ready adjustment to audience preferences when participants with a pro-management attitude had been made accountable to the users of the program than when participants with a pro-user attitude had been made accountable to the management of the program, but again, this effect was not strong enough to be significant.

The analyses performed on the difference between suitability ratings for the chosen vs. average non-chosen alternative indicated that the expected decrease of this difference for accountable participants only happened when they had a pro-user attitude. For participants with a pro-management attitude no significant difference between accountable and non-accountable participants was observed. The reason for this pattern of results was that participants with a pro-management attitude who had been made accountable to the users showed an unexpectedly large evaluative difference between their chosen and average non-chosen alternative, which pulled the mean in the accountability conditions upwards and made it not significantly different from that observed in the 'no accountability' condition. A significant interaction between time of accountability and attitude was obtained when the control group was included in the analysis. Indeed, it was the tendency of non-accountable participants with a pro-user attitude to show a larger judgemental difference between their chosen and average non-chosen alternative than those with a pro-management attitude which caused the interaction effect to become significant. The influence of participants' attitude on this measure can yet again be explained by the fact that for participants with a pro-management attitude, because of their attention to both user and cost attributes, the overall value of the different alternatives they had to choose between was more similar and, therefore, irrespective of their choice, the difference between the chosen and average non-chosen alternative would be expected to be smaller. When participants had been made accountable to either the users of the program or the management of the company, no significant differences between participants with different attitudes emerged. This may be seen as yet another

indication of the fact that participants showed some adjustment to the preferences of their audience and the effect of their own attitude lost its strength.

Contrary to the actually observed evaluative differences between alternatives, the predictions of these differences from participants' unidimensional value and weight judgements, given their more covert nature, had been expected to show evidence of a tendency to defend own preferences rather than adjust to audience preferences when participants experienced a conflict between their own preferences and those of their audience. This was the case for the MAUT predicted evaluative difference between the user and the average management alternative as well as the MAUT predicted evaluative difference between the chosen and the average non-chosen alternative, but, again, usually only when this conflict was one between a pro-user attitude and having been made accountable to the management. Also, the conflict effect tended to be stronger when the change in MAUT predicted evaluative differences from before the first to after the first decision rather than differences obtained after the first decision were analysed.

Interestingly, and in contrast to the results obtained for weight judgements, the MAUT predicted evaluative differences between alternatives showed some effects of the experimental manipulations not only when the predictions were based on the attributes from the first information set, but also when they were based on the parallel attributes from information set 2, although these effects were quite weak and did not usually reach standard significance levels. This may be interpreted as supporting a conclusion that the attempt to make the information contained in the second information set largely redundant to that in the first was successful. Only then would participants have rated attribute outcomes similarly and similar predictions would have been obtained. An important difference to the effects for MAUT predictions based on the attributes from the first information set, however, was that, apart from significant attitude effects, they suggested adjustment to the preferences of the audience participants had been made accountable to rather than conflict effects. This makes sense, given that a conflict is likely to be experienced only when a choice is involved, but participants had not known parallel attribute outcomes for their choice at this point.

Most likely as a result of the differential tendency to adjust to audience preferences by participants with a pro-management and pro-user attitude, respectively, accountable participants did not consistently show an enlarged MAUT predicted difference between the chosen and the average non-chosen alternative compared to participants who had not been made accountable. The only time this effect was found was for the change in the MAUT predicted difference between the chosen and the average non-chosen alternative from before the first to after the first decision, based on the attributes from the second information set. As expected, accountable participants showed a larger increase in this difference than non-accountable participants, who, in fact, showed a decrease, suggesting that regression effects were at work.

An assumption that unfortunately could not be tested adequately, due to the insufficient sample size and asymmetrical choice distributions, was that the way in which a conflict between participants' own attitude and the preferences of the audience they had been made accountable was resolved, that is, either by adjustment to audience preferences or resistance and bolstering of own preferences, would affect the observed evaluative difference between the chosen and average non-chosen alternative. Some tentative analyses, which have to be viewed with caution, given that they were based on very small *n*s, nevertheless suggested that, as expected, when participants decided to resolve the conflict by making a choice according to their own preferences rather than those of the audience they had been made accountable to, the evaluative difference between their chosen and average non-chosen alternative was larger than when they decided to go along with audience preferences.

Information search before the final decision

The defensiveness-inducing effects of resisting audience preferences were supported by the analyses of decision processes observed at the final decision. It had been predicted that participants who had only been made accountable before the final decision would not show adjustment to the preferences of the audience they had been made accountable to, especially if they perceived a conflict between their own preferences and those of the audience they had been made accountable to. This was observed in participants' information search before the

final decision as well as their information evaluation, but, as before, only for participants with a pro-user attitude who had been made accountable to the management.

The defensiveness of these participants' information search, for example, was indicated by the type of information they searched. Participants with a pro-user attitude searched a higher percentage of user information when they had been made accountable to the management than when they had been made accountable to the users of the program, most likely because the majority of them had chosen a user alternative before and needed to defend it. Participants with a pro-management attitude who had been made accountable to the users of the program, on the other hand, searched the same relative amount of user information as participants with a pro-user attitude, suggesting that they adjusted to the preferences of the audience they had been made accountable to.

Furthermore, analyses of the Biased Search Index (BSI) data suggested that the tendency to search more supporting than non-supporting information for the chosen alternative and to search information in such a way that the chosen alternative was supported over the average non-chosen alternative, was stronger for participants with a pro-user attitude than for participants with a pro-management attitude, but only when they had been made accountable before the final decision. This again indicates that the value conflict between a pro-user attitude and having been made accountable to the management was experienced as more severe than the opposite combination, and was particularly strong after a previous commitment.

Unlike expected, accountable participants did not show a significantly stronger bias to search information that supported the chosen alternative over the non-chosen alternatives than participants who had not been made accountable. The analysis performed on the BSI for the average non-chosen alternative even indicated that, compared to non-accountable participants, accountable participants tended to search more information that supported than did not support the non-chosen alternatives. This confirms previously reported findings, for example, by Canon (1964), which suggested that an expectation of having to convince someone else of one's choice results in more attention to dissonant information. It seems reasonable to assume that accountable participants in this study expected to be presented with counter-arguments in

their discussion with the other committee members, particularly when they had chosen to resist the choice that they thought would be favoured by the group they had been made accountable to. They seem to have dealt with this pressure by searching information that could feed into possible counter-arguments, but also made sure that they had more good arguments supporting their own choice.

In addition to any biasing effects, the PMA had also predicted that participants who had been made accountable before the final decision would show a deeper and more complex information processing than participants who had already been made accountable before the first decision or had not been made accountable at all, resulting in significant effects of time of accountability on the information search measures. Accordingly, accountable participants were predicted to show deeper and more complex information processing than non-accountable participants.

The results obtained for the final information search measures largely supported these predictions. As expected, the information search of accountable participants was deeper, more compensatory and less strongly concentrated on the chosen alternative than that of non-accountable participants. Most of the variability of search measures showed a trend that was consistent with these findings, namely, variability of search was lower in the 'accountability' than in the 'no accountability' conditions. An exception was variability of search across attributes, based on the total number of items searched, which tended to be higher for accountable than non-accountable participants. This suggested that accountable participants, when they searched information repeatedly, tended to do so for some alternatives on some attributes to a larger extent than non-accountable participants, possibly to remind themselves of the advantages of the alternative they favoured. When variability of search across attributes was based on different cells searched, however, it indeed tended to be lower when participants had been made accountable before the final decision than when they had either not been made accountable or had been made accountable before the first decision. This shows that participants who had been made accountable before the final decision searched a more equal number of alternatives for each attribute and suggests that they did not eliminate alternatives from consideration at an early stage. The observed variability of search across alternatives was generally larger than variability of search across attributes when both were based on the

different cells searched. This is consistent with a tendency to only search a subset of the available alternatives, but searching a relatively equal number of attributes per considered alternative. Given that participants had eliminated alternatives during their previous choice process already, a strategy that focused on the previously preferred alternative and its closest competitor would make sense.

The expected difference between accountability before the final and accountability before the first decision was significant for some measures, such as concentration of search, but tended to be smaller overall and was sometimes dependent on participants' attitude. For example, a significantly deeper information search for participants who had been made accountable before the final compared to the first decision was only observed for participants with a pro-user attitude. This was caused by the particularly low amount of information that was searched by participants with a pro-management attitude who had been made accountable to the management before the final decision. As a result, the only conditions that showed a significantly deeper information search than the 'no accountability' conditions were the 'accountability to users' conditions. These findings were mirrored in those observed for variability of search across alternatives and complexity of processing. When participants had a pro-user attitude, the observed variability of search was lower and complexity of processing higher when they had been made accountable before the final decision than when they had been made accountable before the first decision, whereas for participants with a pro-management attitude a reverse, but non-significant, tendency was observed. The latter supports the conclusion that participants with a pro-management attitude did not show their usual relatively balanced search of all alternatives when they had been made accountable before the final decision, especially when they had been made accountable to the management. One can only speculate as to why this effect occurred; it may be that because their own preferences did not clash with those of the audience they had been made accountable to, participants felt little need to check upon their previous choice.

Final information evaluation and integration

Participants' final choices were consistent with the adjustment effects predicted by the PMA. When participants had been made accountable only after they had already made a commitment to one of the alternatives, no adjustment was observed and their choices matched their attitude, that is, showed a preference for the user alternative when participants had a pro-user attitude and a preference for management alternatives when participants had a pro-management attitude. When participants had already been made accountable before the first decision, on the other hand, their choices matched the preferences of their external audiences, that is, expressed a preference for the user alternative when they had been made accountable to the users of the program and a preference for the management alternative when they had been made accountable to the management of the company. Somewhat unexpectedly, when participants had not been made accountable, they were more likely to choose the user alternative, even if their own attitude was management-oriented. This was, because some participants changed their initial choice from a management to a user choice, and shows yet again that a user perspective was very strong within the sample.

Participants' weight judgements showed similar effects, but only for the attributes contained in the first information set. As predicted, when participants had been made accountable before the first decision, the mean sum of user attribute weights elicited after the final decision was significantly larger when they had been made accountable to the users of the program than when they had been made accountable to the management of the company, but when they had been made accountable before the final decision, the observed mean sum of user attribute weights did not differ significantly in the two conditions. Moreover, when participants had been made accountable to the management before the final decision, the sum of user attribute weight was larger than when they had been made accountable before the first decision. This is likely to have been the result of participants' initial and final choices, which had been choices of the user alternative in the majority of cases, and which were supported by stressing the importance of user attributes. The only significant effect observed for the weight judgements based on the parallel attributes from information set 2 was one of attitude; as expected, participants with a pro-user attitude assigned higher user attribute weights than

participants with a pro-management attitude. The presence of an effect for the attributes from the first information set but absence of the same effect for the attributes from the second information set, suggests that participants consolidated their choice through their weight judgements for the attributes from information set 1, but did not show differentiation for the new attributes. A possible explanation for this could be the fact that weights for the attributes from the first information set were always elicited first and that, by the time weights for the parallel attributes were elicited, a sufficient bolstering of participants' choice had already occurred. A similar argument has been made by Götz-Marchand, Götz, and Irle (1974), who found empirical support for it in a study where the opportunity to reduce dissonance by negatively evaluating an intelligence test in which participants had allegedly performed badly, was used more often when it was the first dependent measure compared to when it appeared in fifth position in the post-test (see Chapter 4). The fact that no significant effects were obtained either for the *changes* in weight judgements from the second elicitation after the first decision to the third elicitation after the final decision signals that regression effects may have counteracted any importance differentiation and consolidation effects.

A reluctance to adjust to audience preferences when having been made accountable after a commitment to a different alternative was also evident in the actually observed and MAUT predicted evaluative differences between the user alternative and average management alternative, and chosen and average non-chosen alternative, respectively. Furthermore, the observed pattern of results lent again support to the special significance of the conflict between a pro-user attitude and accountability to the management in inducing a particularly strong need to uphold and defend own preferences. It is important to stress that in all instances where this was observed, participants' particular choices could not provide the sole explanation for the effect, because in both conflict conditions (pro-user attitude/accountable to management and pro-management attitude/ accountable to users) the same overall choice distribution was observed; 9 out of 15 participants chose the user alternative.

Actually observed and MAUT predicted evaluative differences between the user and average management alternative were larger when participants had been made accountable to the users before the first rather than the final decision, and larger when they had been made accountable to the management before the final rather than the first decision. This is

consistent with a failure to adjust to audience preferences and bolstering of own preferences when being made accountable before the final decision. In addition, the fact that participants with a pro-user attitude did not show significantly smaller evaluative differences between the user and average management alternative when they had been made accountable to the management rather than the users, whereas participants with a pro-management attitude did, suggests that participants with a pro-user attitude must have been more unwilling to adjust to the preferences of a management audience.

This time, significant effects of a value conflict between a pro-user attitude and accountability to the management were also obtained when changes in the MAUT predicted differences between the user and average management alternative from the first to the final decision were analysed. Unexpectedly, however, this was only the case for predictions based on the attributes contained in the second information set and both information sets combined. This was particularly surprising, given that attribute weights, as important components of these predictions, had only revealed significant effects when the attributes had been contained in information set 1. Another peculiarity of the changes in MAUT predicted differences between the user and average management alternative from the first to the final decision was that time of accountability never significantly interacted with the other factors, suggesting that such effects were not strong enough and/or the measure not sensitive enough to pick them up.

Evaluative differences between the chosen and average non-chosen alternative, both those actually observed and those predicted by an additive linear MAUT model, were found to be increased when participants experienced a conflict between their own attitude and the preferences of the audience they had been made accountable to, particularly when accountability had been introduced before the final rather than the first decision, and when the conflict was between a pro-user attitude and having been made accountable to the management. The triple interaction between all three experimental variables observed in the change of the judgmental difference between the chosen and average non-chosen alternative from the first to the final decision is particularly interesting, because it suggests that a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in an *increased* consolidation when they had been made accountable before the final decision, but in a *decreased* consolidation when they had already been made

accountable before the first decision. This seems to speak against the validity of Svenson's (1992, 1996) hypothesis that individuals will show consolidation even if their choice is not under threat, but seems to support previous findings, for example, by Boiney, Kennedy and Nye (1997), which suggest that participants only bias their information evaluation when this is necessary to arrive at the desired conclusion. Participants with a value conflict between a pro-user attitude and accountability to the management who had already been made accountable before the first decision presumably had already achieved the desired degree of differentiation and therefore did not need to increase the evaluative difference between their chosen and average non-chosen alternative even further.

However, MAUT predicted evaluative differences between the chosen and average non-chosen alternative after the final decision, based on the attributes from the first information set, provided some support for Svenson's theory. A conflict between a pro-user attitude and accountability to the management resulted in a significantly larger evaluative difference than no conflict, whereas the conflict between a pro-management attitude and accountability to the users, although showing the same tendency, did not do so significantly. Although the triple interaction between the experimental factors was not significant, simple main effects analyses suggested that the above was only true for participants who had been made accountable before the first decision. When participants had been made accountable before the final decision, although the same trends were evident, the observed differences were not significant. This points to continued bolstering of the chosen alternative under conflict, even when accountability pressures were not immediately salient anymore.

What could be the reason for these seemingly contradictory results? When participants rated the suitability of the different alternatives, they may have been reluctant to openly show consolidation by rating the overall value of their preferred alternative higher and/or the overall value of their non-chosen alternatives lower than before unless they felt a strong pressure to do so. This may have happened, because they were able to remember their previous judgements and did not want to appear inconsistent. When participants gave unidimensional value and weight judgements, on the other hand, the process of consolidation was much less obvious and previous judgements more difficult to remember, so participants may therefore have engaged in it more readily. The fact that when MAUT predictions were based on the

attributes from the first information set, the consolidation effect was stronger for participants who had been made accountable before the first compared to the final decision may be seen to suggest that accountability before the final decision interfered with consolidation attempts, because it forced participants to reconsider their choice.

Both actually observed and MAUT predicted evaluative differences between the chosen and average non-chosen alternative showed significant effects of a value conflict between participants' own attitude and the preferences of the audience they had been made accountable to when their change from the first to the final decision was analysed. However, when the differences after the final decision, were inspected, significant conflict effects for MAUT predicted differences were only obtained when predictions were based on the first set of attributes or both attribute sets combined rather than based on the second attribute set. The failure to observe any conflict effects for predictions based on the attributes from information set 2 may be explained by the absence of such effects in the weight judgements that participants had given and hence be regarded as a consequence of the particular order in which weight judgements were elicited.

As expected, accountability compared to no accountability resulted in a decreased judgmental difference between the chosen and average non-chosen alternative after the final decision. When inspecting the change of this difference from the first to the final decision, however, this effect was not significant. The MAUT predicted evaluative differences between the chosen and average non-chosen alternative also mostly failed to show the larger means that had been predicted for accountable participants. The only significant effect of accountability vs. no accountability was observed for the change in MAUT predicted evaluative differences between the chosen and average non-chosen alternative from the first to the final decision, based on the attributes from the second information set. The direction of this effect was opposite, however, to what had been expected; accountable participants showed a smaller increase in the MAUT predicted evaluative difference between their chosen and average non-chosen alternative than participants who had not been made accountable.

Justifications

Overall, participants' written justifications were found to focus almost exclusively on the advantages of the chosen alternative and to exhibit a very low complexity, that is, a very low ratio of two-sided arguments (arguments that compared alternatives in some way) to the total number of one- and two-sided arguments. This is consistent with previous findings, for example, by Crozier (1989, 1990), Ranyard (1991) and Huber and Seiser (2001), which showed that justification texts typically only refer to the chosen alternative, which is described very positively. The results of the present data analyses furthermore suggested that a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in a particularly low complexity of argumentation, but only when participants had been made accountable before the first rather than the final decision. This supports the previously reported tendency evident in participants' information evaluation, namely that consolidation was stronger for participants who had been made accountable before the first compared to the final decision. The increased depth and complexity of information processing induced by accountability immediately prior to the final decision seems to have carried over to the justifications.

The effect of conflict was also evident in the relative number of arguments mentioning positive aspects of the chosen alternative, which was larger when participants experienced a conflict between their own preferences and those of the audience they had been made accountable to than when they did not. However, at the same time, the other type of consonant arguments, namely negative aspects of the non-chosen alternatives, was mentioned relatively *less* often when participants experienced a value conflict than when they did not. The reason underlying this finding can only be speculated about. It may be that participants in the conflict conditions who had adjusted their choices to the preferences of the audience they had been made accountable to, particularly avoided mentioning negative aspects of their secretly preferred alternative. Dissonant arguments, that is, arguments that mentioned negative aspects of the chosen alternative or positive aspects of the non-chosen alternatives did not reveal any significant differences between conditions.

Again, results suggested that the conflict between having a pro-user attitude and having been made accountable to the management had a special role in inducing defensive effects, especially when it occurred under post-decisional accountability. Participants who had been made accountable to the management before the final decision elaborated relatively more often on the information that had been presented during the experiment than participants who had been made accountable to the users. Given that these participants had chosen the user alternative in the majority of cases and the elaborations they presented tended to support the user alternative, they seem to have bolstered their choice by means of elaborations.

Participants with a pro-management attitude who had been made accountable to the users, on the other hand, gave incorrect information relatively more often than participants with a pro-management attitude who had been made accountable to the management. The false information they gave did not support the management alternative but the user alternative. This signals adjustment to the preferences of their audience and possibly the wish to make a good impression. In order to achieve this goal, they may have exaggerated the advantages of the user alternative. Participants with the opposite value conflict did not show an equivalent tendency.

Whereas a conflict between a pro-management attitude and having been made accountable to the users of the program resulted in more false reproductions, the absence of this conflict, because participants had a pro-user attitude, seemed to have induced a different strategy, namely a strategy to mention more global evaluation criteria which did not refer to any alternatives in particular, but stressed the importance of user concerns. Participants with a pro-management attitude who had been made accountable to the management, did not show a comparable tendency, which may be explained by the fact that they did not exclusively favour cost concerns over user concerns, and tended to choose compromise alternatives that did not have clear advantages on one or the other.

The PMA had predicted that external accountability to different interest groups would make external norms salient. This should have also found expression in participants' justifications. Indeed, participants justifications showed that they were sensitive to these norms, as participants who had been made accountable to the users of the program mentioned

relatively more user and fewer cost arguments than participants who had been made accountable to the management of the company and vice versa.

It had also been expected that participants who had not been made accountable at any time would exhibit a lower complexity of argumentation, and mention consonant arguments (positive aspects of the chosen alternative and negative aspects of the non-chosen alternatives) relatively more often and dissonant arguments (negative aspects of the chosen alternative and positive aspects of the non-chosen alternatives) relatively less often than accountable participants. Although the observed means pointed in the predicted directions, none of the observed differences were significant. This was also true for the total number of arguments presented, which, as expected, was slightly higher for accountable than for non-accountable participants.

Conclusions

The results of the present study, notwithstanding certain peculiarities resulting from sample characteristics and the particular decision scenario used, supported the predictions of the PMA. External accountability to different interest groups resulted in both complex and biased information processing, no matter whether it was pre-decisional or post-decisional. The fact that when accountability was introduced, information processing remained relatively more extensive and complex, even if participants had already made a choice, suggests that Tetlock's (1985, 1991, 1992) hypothesis that accountability will result in a more thorough information processing only when decision makers have been made accountable before their decision to an audience with unknown views is not quite correct. Also, in line with the predictions derived from the PMA, only pre-decisional accountability induced any significant tendencies to adjust to the preferences of the audience participants had been made accountable to. Post-decisional accountability resulted in defensive reactions, where individuals tried to bolster their initial choice, both by the type of information they searched and the way in which they evaluated and integrated information.

Participants' own preferences appeared as an important moderator of the effects of external accountability on decision processes in this study. A conflict between individuals'

own preferences and those of the audience they were made accountable to typically resulted in defensive information processing that supported participants' own preferences, particularly when accountability was post-decisional. It must be noted, however, that the effect of participants' own preferences may have been overestimated in this study, because unidimensional value and weights judgements were elicited first and therefore may have had an effect similar to a commitment for participants, who may have been concerned to appear consistent afterwards. Further studies need to address this point. Another reason for why the effects of participants' own preferences were particularly pronounced may have been that the consequences of not adjusting to audience preferences but insisting on one's own opinion were, of course, negligible in this artificial laboratory situation. The present results, however, imply that when decisions require adjustment to a particular norm, as it is, for example, the case when individuals act as agents for a principal who expects them to make decisions that match his or her own preferences, they should only be delegated to individuals with similar preferences.

An effect that was not consistently observed, especially where the degree of bias in information evaluation and integration was concerned, was a difference between accountable and non-accountable participants' information processing. Although the direction of observed effects usually matched predictions, they were often not significant. One possible reason for the failure to obtain significant differences is that the accountability pressures created in the present experiment lacked ecological validity and therefore were not strong enough. However, the fact that comparisons between experimental groups usually revealed significant effects seems to speak against this assumption. Instead, task factors seem to have sometimes overshadowed any effects of accountability. For example, the biasing effect of a previous commitment on information search processes was observed in the 'accountability' as well as 'no accountability' conditions. Both accountable and non-accountable participants searched more information for the chosen alternative that supported rather than did not support it, and searched information in such a way that the chosen alternative was supported over the non-chosen alternatives. However, accountable participants also tended to search more information for the non-chosen alternatives that supported rather than did not support these alternatives, whereas non-accountable participants did not show this tendency. This result is consistent

with previous findings which suggest that the expectation of having to convince someone of one's choice increases attention to dissonant information.

Finally, the present study provided some support for Svenson's (1992, 1996) Diff Con theory, in that participants gave unidimensional value and weights judgements which differentiated the chosen alternative from the non-chosen alternatives, particularly when there was a conflict between participants' own preferences and those of the group they had been made accountable to. However, regression effects counteracted these effects and meant that there was not usually a significant *increase* in the evaluative differences between the chosen and average non-chosen alternative, unless a value conflict existed. There was also evidence to support Svenson's assumption of ongoing consolidation after a decision has been made. This tendency was, however, only clearly visible when the measure was a covert one, otherwise there was a suggestion that consolidation only occurred when it was necessary, for example, because participants were under pressure. Hence, like in the previous study, a dissociation between overt and covert measures of differentiation and consolidation was observed. However, this dissociation did not occur when participants experienced a value conflict. In this case, both overt and covert measures showed an enlarged evaluative difference between the chosen and average non-chosen alternative, supporting the assumption that differentiation and consolidation is particularly strong when individuals are under pressure.

The following, final chapter will summarise the main findings from both studies and discuss their implications for theories of the effects of accountability on decision processes in particular and motivation more generally.

CHAPTER 8: GENERAL DISCUSSION AND CONCLUSIONS

The work presented here investigated the effects of accountability on information search, evaluation and integration in multi-attribute decision making, integrating classical social psychological with decision-analytical approaches. Accountability is a social context variable that affects decision processes, like other context variables, such as decision complexity or time pressure. Here it was conceptualised as a motivating agent, that is, as a variable that introduces particular goals into decision makers' information processing. The exact nature of these goals was argued to depend on the particular type of accountability that is created. Contrary to common conceptions of accountability as a unidimensional variable which always produces the same effects on decision processes, a closer inspection of the accountability concept resulted in the conclusion that accountability has two core components; it introduces in the individual a feeling of personal responsibility for potentially negative consequences of his or her behaviour, and it creates evaluation apprehension. Personal responsibility is suggested to the individual by implications of personal causality, that is, personal agency in bringing about certain consequences, and by implications of the individual having control over the consequences of his or her behaviour. Evaluation apprehension, on the other hand, is induced by the implied presence of others and the assumption that these others can identify the individual's performance. While personal responsibility cues make internal standards and norms of behaviour salient, evaluation cues enhance the salience of the norms and standards of the external audience. It was furthermore argued that not all types of accountability contain both components to the same extent. Internal accountability, for example, that is, a need to be able to justify one's behaviour to oneself (Simonson, 1989), suggests personal responsibility to the individual but does not create any significant evaluation apprehension, due to the absence of an external audience. External accountability, on the other hand, implies both personal responsibility and evaluation by an external audience.

A new process model of accountability (PMA, see Chapter 5) was introduced, which linked the antecedents of accountability, personal responsibility and evaluation apprehension, to particular goals with particular cognitive consequences. The model's main premise is that

accountability can make the decision process *both* more complex and more biased and therefore has not the exclusively beneficial effects that it is commonly afforded. Ample evidence from the research literature was presented which supported this claim. The PMA proposes that the increased depth of information search and complexity of information integration that can be observed when individuals are made accountable, is the result of strong accuracy goals, which have been enhanced by the personal responsibility cues inherent in accountability manipulations. The stronger degree of defensiveness and bias in individuals' information evaluation and integration when they have been made accountable, on the other hand, is assumed to be the consequence of directional goals, which are enhanced by the evaluation cues inherent in accountability manipulations. Directional goals, as suggested by their name, are goals that steer the decision process into a particular direction, for example, towards supporting an alternative that is assumed to be favoured by the external audience individuals are accountable to. A directional goal may also be introduced by a previous commitment to an alternative (in this case the decision process is biased towards the previously chosen alternative) or by the individuals' knowledge that the outcome of their decision rather than the process by which they arrived at it will be evaluated[1]. Hence, the different goals that are made salient by personal responsibility and evaluation apprehension, respectively, set into motion different cognitive mechanisms that operate at different stages of the decision process. Accuracy goals enhance depth of information search and complexity of information integration, whereas directional goals introduce a bias into the decision maker's information evaluation and integration (see also Baumeister and Newman, 1994[2]). It is important to note that both tendencies can be compatible when no normatively correct answer to a decision problem exists, as an increased depth of processing may be used to bias the decision process towards a particular outcome and construct justifications for one's conclusions (Kunda, 1990). If a normatively correct answer to a decision problem exists and

[1] The distinction between accuracy and directional goals was originally proposed by Kunda (1990) and can be found in a similar form in Baumeister and Newman's (1994) model of self-regulated inference and decision processes. Baumeister and Newman (1994) distinguish between the mindset of an intuitive scientist, with a need to arrive at an optimal conclusion, and the mindset of an intuitive lawyer, with a need to arrive at a particular conclusion.

[2] Intuitive scientists are assumed to try to achieve their goal primarily by intensifying their information search, intuitive lawyers, on the other hand, are assumed to try to achieve their goal primarily by biasing their information evaluation and integration.

accuracy and directional goals conflict, accountability effects will ultimately depend on which goal is stronger. The question of whether or not accountability makes decision processes more normative as such was not one that this work attempted to answer, but we will nevertheless return to it later.

The PMA predicts that the personal responsibility cues inherent in internal accountability manipulations will induce a deeper information search and more complex information integration compared to no accountability, while any biasing effects on information evaluation and integration processes will be minimal, because of the absence of any evaluation cues. This will, however, only be true as long as internal accountability is introduced before a decision is made. When individuals are made internally accountable after they have already made a decision, although they will still show a deeper information search and more complex information integration compared to non-accountable individuals, they will bias their information evaluation and integration towards supporting the previously chosen alternative. External accountability, on the other hand, will make information processing both more extensive and integratively complex and more biased compared to no accountability, no matter whether it is pre-decisional or post-decisional. The direction of bias induced by post-decisional external accountability, however, is likely to differ from that of pre-decisional external accountability. It is predicted to support the previously chosen alternative, whereas any pre-decisional bias is directed at supporting the alternative that is likely to be favoured by the audience the individual has been made accountable to. In the absence of any salient external norms that can guide the decision maker's information processing, the alternative that appears most defensible, for example, because it has the best outcome on the most important attribute, or because it seems to be a good compromise, is likely to be chosen.

Two empirical studies were conducted to test the above predictions. The first compared internal and external accountability, which was introduced either before or after an initial decision had been made, in order to test whether it would be particularly pre-decisional internal accountability that resulted in an extensive but relatively unbiased decision process. No norms were made explicit in this study. The second study focused on the impact of pre- and post-decisional external accountability to different interest groups, in order to investigate how different external norms would pull the decision process into different directions. In both

studies, participants were asked to search information and make an initial decision before they were given the opportunity to search further information and had to make a final decision. The information offered the second time was largely redundant with that participants were able to search before their first decision, in order to be able to test to what extent information search would be biased to support a previously chosen alternative. Also, unidimensional value and weight judgements were elicited from participants at several points during the experiment, in order to test predictions derived from Svenson's (1992, 1996) differentiation and consolidation theory, particularly attractiveness differentiation and importance differentiation. Both are assumed to be employed in order to increase the overall evaluative difference between the chosen alternative and the non-chosen alternatives. Attractiveness differentiation refers to a change in the evaluation of attribute outcomes, such that outcomes of the chosen alternative are evaluated more positively and/or outcomes of the non-chosen alternatives more negatively than before. Importance differentiation means that the importance of attributes on which the chosen alternative fares well is enhanced and/or the importance of attributes on which the chosen alternative fares badly is decreased. These processes were expected to be more pronounced, when participants were accountable than when they were not accountable.

The effects of accountability on directional bias in information evaluation and integration

Overall, the results supported the predictions of the PMA, particularly in terms of the special role that norms play for directing decision processes. This was suggested by the fact that in the first study, where no particular norms were made salient before participants first made a decision and participants were therefore left to guess the preferences of the audience they had been made accountable to, the effects of a biased information evaluation and integration were rather weak and only marginally significant, whereas in the second study, where the norms of the audience participants were made accountable to were made more explicit, clear biasing effects were observed. This suggests that when individuals are made accountable, impression management concerns become salient and, in order to protect their social image, individuals adapt their expressed preferences to those of the audience they have been made accountable to. These effects were strong enough to find significant expression in

choices and attribute weights and to a lesser extent in suitability ratings of alternatives. Only when participants had not been made accountable, did their choices and attribute importance ratings follow their own attitude. Otherwise, when participants were made accountable to an audience that they expected to have certain preferences, they chose alternatives more often that had positive rather than negative outcomes on attributes participants expected to be considered important by the audience they had been made accountable to, and judged the relative importance of these attributes higher than the importance of attributes that were not expected to be considered important. As predicted, this was only true, however, when participants had been made accountable before they made an initial decision. If accountability was only introduced after participants had already committed themselves to an alternative, participants did not adjust their choice to the preferences of their audience (unless these preferences coincided with their own, obviously) and assigned attribute weights that were consistent with their previous choice.

Although participants were found to rate relative attribute importance in such a way that their chosen alternative was supported, the *change* in these ratings from before the first to after the first decision and after the first to after the final decision was not usually significant, because of relatively strong regression effects, which counteracted differentiation and consolidation effects, particularly in those conditions, in which attribute weights had been extreme. Hence, importance differentiation in the strict sense of the word was not observed.

The effects of attractiveness differentiation were tested by feeding participants' unidimensional value and weight judgements into a linear additive MAUT (multiattribute utility theory) model, which predicted the overall evaluation of alternatives. The reason for why attractiveness differentiation was not analysed independently of importance differentiation is that it is really only effective when it is performed on attributes that are considered important. The obtained results suggested that individuals indeed rate the outcomes of alternatives in such a way that their chosen alternative is supported, creating a large evaluative difference between their chosen and average non-chosen alternative. However, the effects of the experimental variables on the change of these ratings over time were relatively weak and only significant when individuals perceived a conflict between their own preferences and those of the audience they had been made accountable to.

Individuals' own preferences as a moderator of directional effects

The results of the empirical work carried out to investigate the PMA consistently demonstrated the importance of participants' own preferences as a moderator of accountability effects. Any biasing tendencies observed in information evaluation and integration processes were particularly strong when a conflict between participants' own preferences and those of the audience they had been made accountable to existed, especially if participants had already committed themselves to an alternative before they were made accountable and therefore were under particular pressure. In some cases, as mentioned above, defensive bolstering was only observed under such circumstances. There are, however, two reasons why the work presented here may have overestimated the impact of participants' own attitude. The first is a purely methodological reason. The fact that unidimensional values and weights were elicited at the beginning of the experiment may have made participants' own preferences particularly salient to them, and they may have wanted to avoid appearing fickle to the experimenter. The second reason lies in the relative artificiality of the laboratory context participants were exposed to. Participants were only role-playing, had no prior relationship to their audience and could be relatively certain that a failure to adopt the preferences of the audience they had been made accountable to would not result in any severe negative consequences that may otherwise be expected in real-life situations. However, there is some evidence that conflict effects similar to those obtained in the present research can also be observed in the field. Kroon, van Kreveld and Rabbie (1991), for example, in a study on the effects of accountability conducted at two departments of the Dutch Riot Police Academy, found that incongruence between an anti-violence norm introduced by superiors and the norm already existing within the institution resulted in reactance.

Internal vs. external accountability

The effects of pre-decisional internal accountability had been predicted to differ from those of pre-decisional external accountability such that the bias evident in information evaluation and integration would be weaker for participants who had been made internally accountable than for participants who had been made externally accountable. This prediction could not be confirmed, however. If anything, there was a suggestion that participants under

external accountability kept the evaluative difference between their chosen and average non-chosen alternative particularly small. This may partly be explained by the fact that no particular norms had been made salient and, yet again, underlines the significance of norms for directing the decision process. Under normative ambiguity, participants seem to have engaged in a 'sitting on the fence strategy' that would allow them to change their mind without losing face, should it be necessary, and signal recognition of the fact that others may have different preferences. When decision makers were made accountable only after they had already made an initial decision, the results, as predicted, suggested that the information processing of both internally and externally accountable participants became biased to support the previously chosen alternative, but this tendency was stronger for participants under internal than external accountability. This was evident in a number of information search as well as information evaluation measures. For example, participants who had been made internally accountable before the final decision displayed the least compensatory information search and a search that was most strongly concentrated on the chosen alternative. Also, their information search was most strongly biased to search dissonant rather than consonant information for the competing alternative. Interestingly, it was participants who had been made internally accountable before the first decision who showed the least biased information processing when they made their final decision. They also displayed the smallest degree of differentiation between their chosen and average non-chosen as well as directly competing alternative. Whereas for participants who had been made internally accountable before the first decision accuracy goals still seemed to be strongly salient and a motivation to support their previously chosen alternative relatively low, those who had been made internally accountable before the final decision, because of the immediacy of this manipulation, seem to have had a stronger motivation to protect their initial choice, and at the same time lacked the motivation to prepare for a public defence induced by external accountability, which gave them the freedom to show the strongest bias in their information processing.

How much more bias does accountability induce compared to no accountability?

A question which the present work also attempted to answer is how much *more* biased decision processes are under accountability compared to no accountability. Overall, the observed differences between the decision processes of accountable and non-accountable

participants were weaker where the extent of bias in information evaluation and integration was concerned than where depth of information search and complexity of its integration was concerned, and not always significant. This could have partly been the result of accountable participants' tendency not to openly show any differentiation and consolidation processes that were not in accordance with the preferences of the audience they had been made accountable to, which would have meant that any difference between accountable and non-accountable decision makers would become smaller. Evidence which supports this assumption will be discussed in more detail below. However, it may also be the case that a certain degree of bias is normal in choice processes. Indeed, there are theories which argue that decisions are generally made in such a way that they can be defended easily; hence, accountability should not change the choice process dramatically (Tversky, 1972; Slovic, 1975; Slovic, Fischhoff & Lichtenstein, 1988; Shafir, 1993; Shafir, Simonson & Tversky, 1993). For example, Shafir (1993) demonstrated that if individuals are asked to select an alternative, their choice is based on very positive aspects of the alternative, whereas if they are asked to reject an alternative, that is, identify the alternative they would not choose, their rejection is based on very negative aspects of the alternative. This results in the paradox effect that an alternative which has very positive as well as very negative aspects, can both be selected first and rejected first. Preference reversals involving choice and judgement, that is, changes of preference depending on whether individuals have to choose between alternatives or have to rate them in some way, also have been explained with a need to be able to defend choices more than judgements. Slovic, Fischhoff and Lichtenstein (1988) concluded that 'much of the deliberation prior to choice consists of finding a concise, coherent set of reasons that justify the selection of one option over the others' (p. 159).

When is a bias really a bias?

Even if one concludes that accountability has certain effects on information evaluation and integration that are over and above any such effects observed in 'normal', non-accountable choice, the question remains whether one should regard them as biases at all. Many authors, including Lerner and Tetlock (1999), have questioned this terminology. If, like Tetlock (1985, 1991, 1992) proposes, the individual acts like a politician, whose primary goal is to maintain good relationships with his or her constituency, any costs of a bias may be off-set by the

benefits of getting along well with people (Lerner & Tetlock, 1999). Similar arguments have been made with regard to other biases. For example, the compromise effect, that is, the tendency to choose an alternative more often when it becomes a compromise or middle option in a choice set, is rational, if one assumes that the individual attempts to keep blame at a minimal level. Likewise, the fundamental attribution error, that is, the tendency to attribute behaviour to actor rather than situational characteristics, may be an expression of social control, that is, the attempt to hold others strictly accountable for their conduct even if there have been situational constraints on their behaviour (Lerner & Tetlock, 1999). This line of argumentation follows a recent shift towards a pragmatic perspective in social psychology, in the course of which many social psychological phenomena have been re-interpreted in terms of the social utility they have for the individual.

The instrumental nature of directional bias

Apart from one exception, the results did not provide consistent evidence for Svenson's (1992, 1996) assumption that the extent of differentiation and consolidation is independent of the evaluative difference that has been achieved between alternatives already, but rather supported Boiney, Kennedy and Nye's (1997) conclusion that motivated reasoning is instrumental, and that the decision maker biases judgements more or less as needed to support the desired conclusion. This was, for example, indicated by the fact that a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in an *increased* consolidation (evident in an increased difference between participants suitability ratings for their chosen and average non-chosen alternative from the first to the final decision) when they had been made accountable before the final decision, but in a *decreased* consolidation when they had already been made accountable before the first decision.

However, when MAUT predicted rather than actually observed evaluative differences between the chosen and average non-chosen alternatives after the final decision were analysed, there was some suggestion that when these predictions were based on unidimensional value and weight judgements for the attributes contained in the first information search, a conflict between participants' own attitude and the preferences of the audience they had been made accountable to resulted in a significantly enlarged evaluative

difference, particularly for participants who had been made accountable before the first decision. This indicated continued bolstering of the chosen alternative under conflict, even when accountability pressures were not immediately salient anymore.

Methodological boundary conditions of directional bias

This dissociation of results highlighted an important methodological implication that was suggested both by findings from the first and the second study, namely that whether significant differentiation and consolidation effects defending the own choice can be observed or not, depends on whether an overt or a covert measure of differentiation and consolidation is employed. Overt measures, which involve relatively direct comparisons of the quality of the preferred option compared to others, such as actual suitability ratings of alternatives, make it obvious to others when decision makers have biased their information evaluation in order to defend their preferred alternative. Covert measures, on the other hand, such as MAUT predictions, which are calculated from a multitude of single components, hide such processes to a stronger extent. Impression management concerns mean that decision makers do not want to be seen biasing their decision processes too openly, a suggestion that has also been made by authors such as Kunda (1990), Baumeister and Newman (1994) and Hsee (1995, 1996). Hence, it is relatively unlikely that biasing effects will be observed in overt measures[3]. On the contrary, if a danger to be exposed as a non-objective information processor exists, decision makers may keep the evaluative difference between their preferred alternative and other alternatives small, especially if they are made accountable. Study 1 suggested that this tendency to 'sit on the fence' is of particular concern to decision makers when there is normative ambiguity and they cannot be sure about what reaction to expect from their audience. Study 2 suggested that this tendency is also stronger when decision makers anticipate dissent because their choice does not match the preferences of the audience they have been made accountable to. This hypothesis may help to explain the failure of previous studies to find significant differentiation and consolidation effects as a consequence of accountability, and provide a way forward of locating such effects in the future.

[3] Huber and Seiser (2001) indeed did not find any significant difference between accountable and non-accountable participants' suitability ratings for the chosen alternative and for the non-chosen alternatives.

It must also be noted that the fact that in the present studies numerical ratings were elicited rather than ratings on a visual analogue scale, which have been typically employed in previous differentiation and consolidation studies, made it less likely that differentiation and consolidation effects were obtained, because participants were better able to remember their previous judgements, and, for consistency reasons, may have provided these rather than ratings of alternatives, attributes and attribute outcomes as they were perceived at the time of elicitation.

Finally, there was some indirect evidence that the order of dependent measures may have influenced results. For example, in the second study, weights for attributes contained in the first information set were always elicited before weights for attributes contained in the second information set, and it was only for the former that significant differentiation effects were observed. It may be that individuals take the first opportunity to achieve differentiation and consolidation and then do not enhance the evaluative advantage of their chosen alternative any further. This would again confirm the hypothesis that individuals only bias their decision processes as much as is needed, and not more. Support for a similar suggestion in the context of dissonance reduction was provided by Götz-Marchand, Götz, and Irle (1974), who found that the opportunity to reduce dissonance by negatively evaluating an intelligence test in which participants had allegedly performed badly, was preferred when it was the first dependent measure compared to when it was the fifth dependent measure in the post-test.

Another possibility may be that particular ways of achieving differentiation and consolidation, for example, importance differentiation, are preferred over others. Dissonance theory would suggest that such preferences, apart from the extent to which they can hide bias, should depend on the resistance to change of various cognitions. Such effects need to be investigated more systematically.

The effects of accountability on information search

Apart from a biasing effect of salient norms or previous commitments on information evaluation and integration, the PMA had predicted that accountability, irrespective of its type or the time when it was introduced, would make information search more extensive and

information integration more complex, compared to conditions where accountability was absent. Overall, this tended to be the case, both in the first and in the second study, and was evident in a higher depth of search, lower variability of search, higher Index of Compensatory Processing and lower concentration of search on the chosen alternative under accountability, compared no accountability. These findings are consistent with previous research, for example, by Huber and Seiser (2001), who found that participants who had been made accountable before they made a decision searched more information, both overall and repeatedly, compared to participants who had not been made accountable.

The fact that information processing remained at a deeper level, even when participants were only made accountable after they had already made a decision, contradicts Tetlock's (1985, 1991, 1992) hypothesis that the integrative complexity of individuals' information processing is only high when individuals are made accountable to an audience with unknown views before they make a decision, but supports the assumption of the PMA that accountability generally increases depth of processing. The difference between accountability before a first and before a final decision (rather than between accountability and no accountability) was less pronounced, but still approached significance in some cases.

The particular type of accountability participants experienced also affected information search processes, by increasing the attention to particular attributes. For example, participants who had been made accountable to a particular interest group concentrated their information search on those attributes that they expected to be considered important by the group they had been made accountable to. However, as predicted, this was generally only the case if they had been made accountable before the first decision. Participants who had been made accountable before the final decision mainly searched information that was relevant to their previous choice, and therefore did not show a significant adjustment to the preferences of their audience when these conflicted with their own. Also, as expected, when participants had been made accountable to an audience with unknown views, their information search showed little evidence of being focused on particular attributes.

There was also some effect of type of accountability on the extent of bias displayed in participants' final information search. As expected, externally accountable participants showed a significant tendency to search more information that they expected to support rather

than contradict their chosen alternative, and they did significantly more so than internally accountable participants. They also showed a significantly stronger bias to support the chosen over the competing alternative than participants in the 'internal accountability' conditions. Furthermore, participants who experienced a value conflict between their own preferences and those of the audience they had been made accountable to tended to exhibit a stronger bias to support their chosen alternative and to support their chosen over their average non-chosen alternative than did participants without such a conflict.

The effects of task variables on information search

At the same time as being affected by the experimental variables of time and type of accountability, the influence of task variables, such as the complexity of the task, the task of having to judge all alternatives, or a previous commitment was strongly evident in participants' information search. The relatively high informational complexity of both decision tasks meant that non-compensatory processing was common, where decision makers did not trade off attributes, but seemed to have eliminated alternatives from their consideration, if they did not meet certain criteria. In study 1, this effect was counteracted by pre-decisional instructions to rate all alternatives, in addition to making a choice. This meant that participants engaged in a generally very extensive information search. After they had made a decision, participants generally did not search information for all alternatives anymore, but concentrated their search on a sub-set of alternatives. In study 1, where a stalemate between alternatives was suggested, this sub-set typically included the chosen alternative and the other stalemate alternative as its closest competitor. Information search after a previous commitment was also found to be generally biased to support the chosen over the average non-chosen alternative. In the second study, there was also a significant general tendency to search more supporting than non-supporting information for the chosen alternative. These findings illustrate the effects of context variables other than accountability, which sometimes overshadowed accountability effects.

The effects of accountability on justifications

Participants' written justifications of their choices showed some evidence of the higher complexity in information processing induced by accountability, although the effects observed here were not usually strong enough to be significant. Overall, the complexity of argumentation (i.e., the relative frequency of comparison of alternatives) was very low and justifications almost exclusively referred to the advantages of the chosen alternative. However, the justifications of accountable participants tended to be more complex and mentioned consonant arguments relatively less often and dissonant arguments relatively more often than those of accountable participants. Accountable participants also tended to mention more arguments overall than non-accountable participants. At the same time, the justifications of accountable participants were clearly directed toward the preferences of the audience they had been made accountable to, evident in the different relative percentage of arguments focusing on attributes that would have been perceived as relevant by the audience participants had been made accountable to. This supports Kunda's (1990) assumption that extensive information processing may facilitate the construction of justifications for desired conclusions.

The justifications of externally accountable participants were found to contain relatively fewer arguments referring to negative aspects of the chosen, but relatively more arguments referring to positive aspects of the non-chosen alternatives than those of internally accountable participants. This may be regarded as yet another indicator of the heightened impression management concerns under external accountability; externally accountable participants avoided letting their own choice appear vulnerable, but showed that they were aware of the advantages of other alternatives preferred by their external audience.

A tendency to bolster the chosen alternative by mentioning a large number of positive aspects of the chosen alternative was particularly pronounced, when participants perceived a conflict between their own attitude and the preferences of the audience they had been made accountable to. However, the relative number of arguments mentioning negative aspects of the non-chosen alternatives was smaller when a value conflict existed, presumably because some individuals had chosen to give in to audience pressures and did not want to mention negative aspects of their truly preferred alternative. A value conflict furthermore resulted in a

relatively larger percentage of falsely reproduced information, probably indicating a strategic rather than a real failure to remember information correctly. Finally, a value conflict resulted in a lower complexity of argumentation compared to the absence of such a conflict when participants had been made accountable before the first, but not when they had been made accountable before the final decision. Therefore, the higher complexity of processing induced by a more recent anticipation of having to justify one's choice to someone with a different opinion seems to have carried over to the justifications. A strategy of participants whose attitude coincided with that of their audience, on the other hand, seems to have been to mention general evaluation criteria they could expect their audience to share. They were found to do so relatively more often than participants who experienced a value conflict. Huber and Seiser (2001) had found a significant difference between accountable and non-accountable participants in this respect. This finding was not confirmed by the present results, however.

Does accountability improve decision processes?

The decision problems used in the empirical investigations here did not have an obvious normatively correct solution (in fact, alternatives were even constructed in such a way that their overall utility predicted by a MAUT model would be very similar if attributes were weighted equally), hence the obtained findings cannot speak to the question of whether accountability results in *better* decisions. The present findings only imply that decision processes *change* when participants are made accountable. If a normatively correct answer exists, previous research has shown that accountability only increases decision accuracy if participants know or can anticipate this answer, or if the increased depth of processing induced by accountability makes it more likely that the normative answer is identified (Simonson & Nye, 1992). Hence, when participants have been trained in using certain normative decision rules, for example, to ignore sunk costs, the sunk cost effect is attenuated by accountability, and so is the primacy effect, due to a more extensive information processing, but choice-matching preference reversals, for example, are not. The more extensive information processing induced by accountability can backfire, however. Under accountability, the dilution effect, that is, the tendency to inappropriately take into account non-diagnostic information which dilutes the extremity of predictions (Nisbett, Zukier, &

Lemley, 1981; Zukier, 1982), is enhanced (Tetlock & Boettger, 1989, Tetlock, Lerner, & Boettger, 1996).

It may be argued that the existence of an obviously correct solution which, if not known, can be identified through an increased information processing effort, is relatively uncommon in decision making under certainty, and, therefore, the likelihood that accountability can result in an improvement of the accuracy of such decisions is not very high.

Qualitative vs. quantitative motivated reasoning

One of the ongoing controversies in the motivated reasoning literature is whether motivational goals change the nature of information processing, that is, have a qualitative effect, or only change the intensity of information processing, that is, have a quantitative effect. A supporter of the former view is, for example, Kunda (1990, 1999) with her distinction between accuracy and directional goals, which are assumed to be met by different cognitive operations, a view this thesis also has subscribed to. Ditto and his co-workers (Ditto & Lopez, 1992; Ditto, Scepansky, Munro, Apanovitch and Lockhart, 1998), on the other hand, are supporters of the latter view, and suggest that the reason why individuals accept preference-consistent information more readily than preference-inconsistent information is that preference-consistent information is less likely to initiate effortful cognitive analysis than preference-inconsistent information. While there is evidence to support this hypothesis, the present results seem to quite clearly point to two distinct effects of accountability, an increase in depth of information search and complexity of information integration and an increase of directional bias in information evaluation and integration. Analysis of the extent of bias in the final information search furthermore did not suggest that preference-consistent information was processed to a lesser extent. In study 1, externally accountable participants searched more information that supported their chosen alternative over their average non-chosen alternative than internally accountable participants. In study 2, participants who experienced a value conflict between their own preferences and those of the audience they had been made accountable to also showed this tendency. Ditto et al. (1998) concede that any factor that increases either the incentive for accurate judgements or the penalty for inaccurate ones, and

accountability is one such factor, should promote effortful processing of information people want to believe and this seems to have been observed here.

The present findings also underline the validity of Kruglanski's (1996) contention that, in order to be able to predict motivated reasoning effects accurately, it is important to both identify the type of goal, that is, the desired end state, and its strength or magnitude. Whereas the former exerts an influence on the *direction* of cognitive activity, the latter affects the *extent* of cognitive activity. Because goal magnitude quantifies the effects of a goal type, according to Kruglanski, it is impossible to predict specific motivational effects on cognition from magnitude alone. Likewise, because of the principle of equifinality, which suggests that a goal can be reached in different ways, the identification of the goal alone is also insufficient for arriving at correct predictions about particular cognitive effects. Only a careful analysis of the psychological situation the individual is confronted with allows to predict which special means he or she is likely to choose to reach a particular goal, and the vigour with which it will be employed. In the present work, the analysis of the antecedents of different types of accountability and the normative context appeared vital for correctly predicting specific accountability effects.

Accountability effects in natural environments

The fact that significant adjustment to audience norms could be observed, despite the fact that accountability was manipulated in a laboratory context, implies that similar effects should occur when audience pressures are real. The social psychological persuasion and conformity literature would furthermore suggest that audience variables, such as the audience's reward and punitive power and likeability should influence participants' willingness to adjust to audience preferences. Studies by Cvetkovich (1978), Gordon and Stuecher (1992) and Haccoun and Klimoski (1975), for example, have provided some support for these assumptions.

An important difference that a real-world context, such as an organisational environment, may present, however, is familiarity with the task. The participants in the present studies were students, and as such were not particularly experienced with the kind of decision problems they were presented with. Experts may not be as susceptible to the effects

of social pressure, because they have confidence in their own judgement and are used to following certain procedures. However, research carried out by Janis on the groupthink phenomenon (Janis, 1982) has illustrated how even for extremely consequential decisions, such as the US government's decision to invade Cuba in 1961, social pressures overrode expertise and resulted in highly defensive, dysfunctional information processing. Also, the controversy about whether experts are less susceptible to effects of motivated reasoning more generally, and decision biases in particular, is an on-going one. While Smith and Kida (1991), for example, have shown that professional auditors are less prone to biases, the medical decision making literature suggests that medical experts are as much, if not more, liable to certain decision biases as lay people (e.g., Heller, Saltzstein, & Caspe, 1992).

Does accountability change preferences permanently?

Another question that arises is, whether the effects that accountability has on the way in which information is processed, result in a permanent change of individuals' preferences. Given that accountability increases the individual's focus on salient norms, its effect may be expected to be more similar to normative influence, which typically only results in a temporary change of opinion, compared to informational influence, which changes attitudes more permanently (e.g., Deutsch & Gerrard, 1955). Simonson and Nye (1992) found some support for conformity processes as a result of pre-decisional external accountability. In one of their studies, participants who were simply instructed to process information thoroughly did not exhibit the sunk cost effect to a lesser extent than a control group who had not been made accountable. Participants who had been made accountable, on the other hand, showed a reduction in this bias. This implies that it was not the increased depth of processing of information but an adjustment to what accountable participants considered a normative response that reduced this bias. Also, in another study by Irwin and Davies (1995), responses to socially significant decision problems after a group choice returned to their original individual level, suggesting that adjustment to the values of the group was only short-lived. One way to investigate such effects more systematically is to employ an audience cancellation method, where participants first expect to have to justify their decision to some external audience, but then, after they have made a decision, are told that they will not have to do so after all. If their response is different from that of participants who are not given the audience

cancellation instruction, this would suggest that any observed preference shift is only strategic. Pennington and Schlenker's (1999) study, which employed such a paradigm, seemed to imply that external accountability does not only produce normative but also informational influence. Participants in an audience cancellation condition still showed some conformity to the norms of the audience they had been made accountable to, even when they made their decision only after having learned that the expected meeting with their audience had been cancelled, although to a lesser extent. In this case, it could not be ruled out, however, that participants may have believed that a meeting between them and their audience would take place at a later time.

Implications for the use of accountability as a means of control in organisations

Considerations like the ones just discussed have important implications for the use of accountability as a norm-enforcement mechanism in applied contexts. If pre-decisional external accountability induces only temporary adjustments to salient norms, it is only useful if the accountability pressure exerted by principals who delegate decisions to their agents can be maintained salient over a certain time span. This constant monitoring of agents' behaviour is costly, however, and somewhat undermines the goal of reducing workload by delegating decisions in the first place. Also, the larger the organisation, the more difficult it becomes to ensure accountability of subordinates to organisational leaders, and rules and regulations typically replace direct supervision, which are easier to ignore (Jos & Tompkins, 1994). At the same time, subordinates may not even be aware of whom they are accountable to and for what.

Furthermore, some of the results presented here have suggested that if there is a conflict between the decision maker's own preferences and those of the audience he or she is accountable to, information processing may become defensive, with the aim of supporting the decision maker's own preference, particularly if accountability was only introduced after a decision had already been made. The extent of the perceived discrepancy between preferences should moderate this effect, however. If it is only slight to medium, the rewards and/or punishments associated with adjustment and a failure to adjust, respectively, may be strong enough to induce conformity. Indeed, there was some evidence for the truth of this assumption in the present data. In the second study, participants with a pro-management

attitude who had been made accountable to the future users of a program, the purchase of which they were asked to make a decision on, adjusted to the preferences of their audience more readily than participants with a pro-user attitude who had been made accountable to the management.

One way to avoid reactance effects may be to delegate decisions only to agents who share the principal's values. This suggests that when organisations select personnel, such considerations should play a role. Alternatively, the goal of organisational culture should be to encourage employees to internalise the value system of the organisation, so that they can be trusted to make decisions in accordance with that system. The fact that in the present studies pre-decisional internal accountability overall seemed to have the fewest negative effects and kept information processing relatively balanced, even after an initial decision had been made, supports the potential benefits of such an approach. However, it is important to stress that it was only *pre-decisional* internal accountability that had such an effect. Post-decisional internal accountability induced information processing that was strongly biased, even more so than when accountability was external.

Another condition that seemed to have had relatively beneficial effects was normative ambiguity, at least in the sense that decision makers were unable to demonstrate adjustment to certain norms and therefore engaged in a 'sitting on the fence strategy'. However, normative ambiguity can arise in two different ways. Firstly, the individual may genuinely not be able to anticipate the preferences of the audience he or she is accountable to, because there are no cues as to what these preferences are and/or these preferences are not very pronounced. Secondly, and this situation seems to be much more common, the individual can anticipate the preferences of the audience he or she is accountable to, but is equally accountable to more than one person or interest group whose views conflict with each other. This is, for example, the case, when companies are both accountable to their clients and shareholders, or when drug regulation agencies are accountable to both the pharmaceutical industry and patients who take the drugs that are admitted to or removed from the market. Tetlock and Boettger's (1994) investigation of a scenario similar to the latter situation suggested that, when made accountable to multiple audiences with conflicting views, individuals often attempt to cope by buck-passing, that is, trying to transfer responsibility for the decision to others, and by procrastinating, that is, delaying the decision. Hence, the kind of normative ambiguity that is

most common may result in dysfunctional behaviour. It must be mentioned at this point, however, that there are some studies which have suggested that accountability to an audience with conflicting views may enhance complexity of processing (e.g., Carnevale & Mack, reported in Carnevale, 1985), and may induce more constructive negotiation strategies (Pruitt & Carnevale, 1982). However, it looks like such results may only be obtained if compromises are possible. As to the first type of normative ambiguity, further research is necessary to systematically investigate the conditions under which individuals attempt to guess the preferences of their audience, which type of cues they base their guess on, and which type of cues will promote success or failure at guessing correctly. Biases such as the false consensus effect, that is, the tendency to assume that others share one's own opinions (Ross, Greene, & House, 1977), suggest that individuals may display an egocentric bias in their guesses and often guess incorrectly.

The work presented here has clearly demonstrated that accountability is not universally beneficial, due to its potential to focus individuals' attention on salient norms and induce information processing that is directed to favour a conclusion compatible with these norms. Whereas this norm enforcement aspect of accountability is vital for social control and maintaining social order, it can also be dysfunctional. Moreover, a conflict between individuals' preferences and the norms of the audience they have been made accountable to seems to result in particularly defensive information processing. Conditions in which external norms are completely absent are rare and such effects, therefore, are difficult to avoid. Humans as social animals are invariably subject to social influence and this influence is also evident in decision processes.

REFERENCES

Aaker, J. (1991). The negative attraction effect? A study of the attraction effect under judgment and choice. *Advances in Consumer Research*, *18*, 462-469.

Abelson, R. P. (1963). Computer simulation of "hot" cognition. In S. S. Tomkins & S. Messick (Eds.), *Computer simulation of personality* (pp. 277-298). New York: Wiley.

Abelson, R. P., & Levi, A. (1985). Decision making and decision theory. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (Vol. 1, pp. 231-309). New York: Random House.

Adelberg, S., & Batson, C. D. (1978). Accountability and helping: When needs exceed resources. *Journal of Personality and Social Psychology*, *36*, 343-350.

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior experience and current situational information. *Psychological Review, 91*, 112-149.

Anderson, P. A. (1981). Justifications and precedents as constraints in foreign policy decision-making. *American Journal of Political Science*, *25*, 738-761.

Antaki, C. (1988). Explanations, communication, and social cognition. In C. Antaki (Ed.), *Analyzing everyday explanation: A casebook of methods* (pp. 268-282). London: Sage.

Antonioni, D. (1994). The effects of feedback accountability on upward appraisal ratings. *Personnel Psychology, 47*, 349-356.

Arkes, H. R. (1996). The temperature of Diff Con theory. *Organizational Behavior and Human Decision Processes, 65*, 268-271.

Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112*, 117-135.

Aronson, E. (1968). Dissonance theory: Progress and problems. In R. P. Abelson, E. Aronson, W. J. McGuire, T. M. Newcomb, M. J. Rosenberg, & P. H. Tannenbaum (Eds.), *Theories of cognitive consistency: A sourcebook* (pp. 5-27). Chicago: Rand McNally.

Ashton, R. H. (1992). Effects of justification and a mechanical aid on judgment performance. *Organizational Behavior and Human Decision Processes, 52*, 292-306.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215.

Bass, B. M. (1966). Effects on the subsequent performance of negotiations of studying issues or planning strategies alone or in groups. *Psychological Monographs, 80*, 1-31.

Bassok, M., & Trope, Y. (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? *Social Cognition, 2*, 199-216.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*, 3-26.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46*, 610-620.

Baumeister, R. F., & Cairns, K. J. (1992). Repression and self-presentation: When audiences interfere with self-deceptive strategies. *Journal of Personality and Social Psychology, 62*, 851-862.

Baumeister, R. F., & Newman, L. S. (1994). Self-regulation of cognitive inference and decision processes. *Personality and Social Psychological Bulletin, 20* , 3-19.

Bayertz, K. (Ed.) (1996). *Verantwortung - Prinzip oder Problem?* Darmstadt: Wissenschaftliche Buchgesellschaft.

Bazerman, M. H., Tenbrunsel, A. E., & Wade-Benzoni, K. (1998). Negotiating with yourself and losing: Making decisions with competing internal preferences. *Academy of Management Review, 23*, 225-241.

Beach, L. T., & Mitchell, T. R. (1978). A contingency model for selection of decision strategies. *Academy of Management Review, 3*, 439-449.

Bell, D. E., Keeney, R. L., & Raiffa, H. (1977). *Conflicting objectives in decisions*. New York: Wiley.

Benson III, L. (1993). *Studies in human decision making: On the effects of experimental instructions, framing, and time constraints*. Unpublished doctoral dissertation, Lund University.

Benson III, L., & Svenson, O. (1992). Post-decision consolidation following the debriefing of subjects about experimental manipulations affecting their prior decisions. *Psychological Research Bulletin, 32*, Lund University

Ben-Yoav, O., & Pruitt, D. G. (1984). Accountability to constituents: A two-edged sword. *Organizational Behavior and Human Performance, 34*, 283-295.

Benthorn, L. J. (1994). *On post-decision processes*. Unpublished doctoral dissertation, Lund University.

Benton, A. A. (1972). Accountability and negotiations between group representatives. *Proceedings of the 80th Annual Convention of the APA*, 227-228.

Beyth-Marom, R., & Fischhoff, B. (1983). Diagnosticity and pseudodiagnosticity. *Journal of Personality and Social Psychology, 45*, 1185-1195.

Billings, R. S., & Scherer, L. L. (1988). The effects of response mode and importance on decision-making strategies: Judgment versus choice. *Organizational Behavior and Human Decision Processes, 41*, 1-19.

Bodenhausen, G. V., Kramer, G. P., & Süsser, K. (1994). Happiness and stereotypic thinking in social judgement. *Journal of Personality and Social Psychology, 66*, 621-632.

Bond, C. F., Jr. (1982). Social facilitation: A self-presentational view. *Journal of Personality and Social Psychology, 42*, 1042-1050.

Borcherding, K. (1981). *On the successive evaluation of multiattribute decision alternatives*. Unpublished manuscript.

Borcherding, K. (1983). Entscheidungstheorie und Entscheidungshilfeverfahren für komplexe Entscheidungssituationen. In M. Irle (Ed.), *Methoden und Anwendungen in der Marktpsychologie, Enzyklopädie der Psychologie* (Vol. D/III/5, pp. 64-173). Göttingen: Hogrefe.

Borcherding, K. (1991). Die Psychologie des Entscheidens. In H.-J. Gamm (Ed.), *Die Wissenschaft am Ende des 20. Jahrhunderts: Ringvorlesung im WS 1989/1990* (Vol. 54, THD Schriftenreihe Wissenschaft und Technik), Technische Hochschule Darmstadt.

Borcherding, K., Eppel, T., & von Winterfeldt, D. (1991). Comparison of weighting judgments in multiattribute utility measurement. *Management Science, 37*, 1603-1619.

Borcherding, K., Schmeer, S., & Weber, M. (1995). Biases in multiattribute weight elicitation. In J.-P. Caverni, M. Bar-Hillel, F.H. Barron, & H. Jungermann (Eds.), *Contributions to Decision-Making I* (pp. 1-28). Amsterdam: North-Holland.

Borgida, E., & Howard-Pitney, B. (1983). Personal involvement and the robustness of perceptual salience effects. *Journal of Personality and Social Psychology, 45*, 560-570.

Boudreau, L. A., Baron, R. M., & Oliver, P. V. (1992). Effects of expected communication target expertise and timing of set on trait use in person description. *Personality and Social Psychology Bulletin, 18*, 447-451.

Billings, R. S., & Scherer, L. (1988). The effects of response mode and importance on decision-making strategies: Judgment versus choice. *Organizational Behavior and Human Decision Processes, 41*, 1-19.

Boiney, L. G., Kennedy, J., & Nye, P. (1997). Instrumental bias in motivated reasoning: More when more is needed. *Organizational Behavior and Human Decision Processes, 72*, 1-24.

Borgida, E., & Howard-Pitney, B. (1983). Personal involvement and the robustness of perceptual salience effects. *Journal of Personality and Social Psychology, 45*, 560-570.

Bowen, J., & Qiu Z-L. (1992). Satisficing when buying information. *Organizational Behavior and Human Decision Processes, 51*, 471-481.

Brehm, J. (1956). Post-decision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52*, 384-389.

Brehm, J. W., & Cohen, A. R. (1962). *Explorations in cognitive dissonance*. New York: Wiley.

Brief, A. P., Dukerich, J. M., & Doran, L. I. (1991). Resolving ethical dilemmas in management: Experimental investigations of values, accountability, and choice. *Journal of Applied Social Psychology, 21*, 380-396.

Brockner, J., Rubin, J. Z., & Lang, E. (1981). Face-saving and entrapment. *Journal of Experimental Social Psychology, 17*, 68-79.

Brockner, J., Shaw, M. C., & Rubin, J. Z. (1979). Factors affecting withdrawal from an escalating conflict: Quitting before it's too late. *Journal of Experimental Social Psychology, 15*, 492-503.

Buttny, R. (1993). *Social accountability in communication*. London: Sage.

Canon, L. K. (1964). Self-confidence and selective exposure to information. In L. Festinger (Ed.), *Conflict, decision, and dissonance* (pp. 83-95). Stanford, CA: Stanford University Press.

Carnevale, P. J. D. (1985). Accountability of group representatives and intergroup relations. In E. J. Lawler (Ed.), *Advances in group processes* (Vol. 2, pp. 227-248). Greenwich, Connecticut: JAI Press.

Carnevale, P. J. D., Pruitt, D. G., & Seilheimer, S. D. (1981). Looking and competing: Accountability and visual access in integrative bargaining. *Journal of Personality and Social Psychology, 40*, 111-120.

Carver, C. S., & Scheier, M. F. (1981a). *Attention and self-regulation: A control theory approach to human behaviour*. New York: Springer.

Carver, C. S., & Scheier, M. F. (1981b). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology, 17*, 545-568.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*, 752-766.

Chaiken, S., Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 212-252). New York: The Guilford Press.

Chen, S., Shecter, D., & Chaiken, S. (1996). Getting at the truth or getting along: Accuracy-versus impression-motivated heuristic and systematic processing. *Journal of Personality and Social Psychology, 71*, 262-275.

Chinburapa, V., Larson, L. N., Brucks, M., Draugalis, J., Bootman, J. L., & Puto, C.P. (1993). Physician prescribing decisions: The effects of situational involvement and task complexity on information acquisition and decision making. *Social Science & Medicine, 36*, 1473-1482.

Christensen-Szalanski, J. J. J. (1978). Problem solving strategies: A selection mechanism, some implications, and some data. *Organizational Behavior and Human Performance, 22*, 307-323.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 201-234). San Diego: Academic Press.

Clarke, P., & James, J. (1967). The effects of situation, attitude intensity , and personality on information seeking. *Sociometry, 30*, 235-245.

Conlon, E. J., & Wolf, G. (1980). The moderating effects of strategy, visibility, and involvement on allocation behavior: An extension of Staw's escalation paradigm. *Organizational Behavior and Human Performance, 26*, 172-192.

Cooper, J. & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17, pp. 229-265). New York: Academic Press.

Cottrell, N. B. (1972). Social facilitation. In C. G. McClintock (Ed.), *Experimental social psychology* (pp. 185-236). New York: Holt, Rinehart, & Winston.

Cronbach, L. J. (1955). Processes affecting scores on "understanding others" and "assumed similarity". *Psychological Bulletin, 52*, 177-193.

Creyer, E., & Ross, Jr., W. T. (1982). Hindsight bias and inferences in choice: The mediating effect of cognitive effort. *Organizational Behavior and Human Decision Processes, 55*, 61-77.

Crozier, W.R. (1989). Post-decisional justification: The case of De Lorean. In H. Montgomery & O. Svenson (Eds.), *Process and structure in human decision making* (pp. 275-292). Chichester: Wiley.

Crozier, W. R. (1990). Accounting for decisions: Justification and rhetoric. In K. J. Gilhooly, M. T. G. Keane, R. H. Logie, & G. Erdos (Eds.), *Lines of thinking: Reflections on the psychology of thought* (Vol. 1, pp. 315-325). Chichester: Wiley.

Curley, S. P., Yates, J. F., & Abrams, R. A. (1986). Psychological sources of ambiguity avoidance. *Organizational Behavior and Human Decision Processes, 38*, 230-256.

Cvetkovich, G. (1978). Cognitive accommodation, language, and social responsibility. *Social Psychology, 41*, 149-155.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20-33.

Dawes, R. M. (1979). The robust beauty of improper linear models. *American Psychologist, 34*, 571-582.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81*, 95-106.

De Hoog, R., & van der Wittenboer, G. (1986). Decision justification, information structure and the choice of decision rules. In B. Brehmer, H. Jungermann, P. Lourens, & G.

Sevón (Eds.), *New directions in research on decision making* (pp. 191-204). North-Holland: Elsevier.

Deutsch, M., & Gerard, H. G. (1955). A study of normative and informational social influence upon individual judgment. *Journal of Abnormal and Social Psychology, 51*, 629-636.

Devine, P. J. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.

Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and conformation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology, 58*, 952-963.

Ditto, P. H., & Lopez, D. F. (1992). Motivated skepticism: Use of differential decision criteria for preferred and nonpreferred conclusions. *Journal of Personality and Social Psychology, 63*, 568-584.

Ditto, P. H., Scepansky, J. A., Munro, G. D., Apanovitch, A. M., & Lockhart, L. K. (1998). Motivated sensitivity to preference-inconsistent information. *Journal of Personality and Social Psychology, 75*, 53-69.

Doherty, M. E., Mynatt, C. R., Tweney, R. D., & Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychologica, 43*, 111-121.

Doney, P. M., & Armstrong, G. M. (1986). Effects of accountability on symbolic information search and information analysis by organizational buyers. *Journal of the Academy of Marketing Science, 24*, 57-65.

Druckman, D. (1968). Prenegotiation experience and dyadic conflict resolution in a bargaining situation. *Journal of Experimental Social Psychology, 4*, 367-383.

Dunning, D., Perie, M., & Story, A. L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology, 61*, 957-968.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57*, 1082-1090.

Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review, 66*, 183-201.

Edwards, W. (1971). Social utilities. *Engineering Economist, Summer Symposium Series, 6*, 119-129.

Edwards, W. (1977). How to use multiattribute utility measurement for social decision making. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-7*, 326-340.

Eisenhardt, K. M. (1989). Agency theory: An assessment and review. *Academy of Management Review, 14*, 57-74.

Emby, C., & Gibbins, M. (1988). Good judgement in public accounting: Quality and justification. *Contemporary Accounting Research, 4*, 287-313.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences*. London: Lawrence Erlbaum.

Fandt, P. M. (1991). The relationship of accountability and interdependent behavior to enhancing team consequences. *Group & Organization Studies, 16*, 300-312.

Fandt, P. M., & Ferris, G. R. (1990). The management of information and impressions: When employees behave opportunistically. *Organizational Behavior and Human Decision Processes, 45*, 140-158.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Festinger, L. (Ed.) (1964). *Conflict, decision, and dissonance*. Stanford, CA: Stanford University Press.

Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of deindividuation in a group. *Journal of Abnormal and Social Psychology, 47*, 382-389.

Fincham, F. D., & Jaspers, J. M. (1980). Attribution of responsibility: From man the scientist to man as lawyer. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 81-138). New York: Academic Press.

Fischhoff, B. (1982), Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 422-444). Cambridge: Cambridge University Press.

Fischhoff, B., & Beyth-Marom, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90*, 239-260.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. New York: Random House.

Ford, J. K., & Weldon, E. (1981). Forewarning and accountability: Effects on memory-based interpersonal judgments. *Personality and Social Psychology Bulletin, 7*, 264-268.

Foschi, M. (1996). Double standards in the evaluation of men and women. *Social Psychology Quarterly, 59*, 237-254.

Freedman, J. L. (1965a). Preference for dissonant information. *Journal of Personality and Social Psychology, 2*, 287-289.

Freedman, J. L. (1965b). Confidence, utility, and selective exposure: A partial replication. *Journal of Personality and Social Psychology, 2*, 778-780.

Freedman, J. L., & Sears, D. O. (1965). Selective exposure. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 57-97). New York: Academic Press.

Freund, T., Kruglanski, A. W., & Shpitzajzen, A. (1985). The freezing and unfreezing of impressional primacy: Effects of the need for structure and the fear of invalidity. *Personality and Social Psychology Bulletin, 11*, 49-487.

Frey, D. (1978). A clarification of selective exposure. The impact of choice. *Journal of Experimental Social Psychology, 14*, 132-139.

Frey, D. (1981a). Reversible and irreversible decisions: Preference for consonant information as a function of attractiveness of decision alternatives. *Personality and Social Psychology Bulletin, 7*, 621-626.

Frey, D. (1981b). The effect of negative feedback about oneself and cost of information of preferences for information about the source of this feedback. *Journal of Experimental Social Psychology, 17*, 42-50.

Frey, D. (1981c). Postdecisional preference for decision relevant information as a function of the competence of its source and the degree of familiarity with this. *Journal of Experimental Social Psychology, 17*, 51-67.

Frey, D. (1982). Different levels of cognitive dissonance, information seeking, and information avoidance. *Journal of Personality and Social Psychology, 43*, 1175-1183.

Frey, D. (1985). *Amount of available information and selective exposure*. Unpublished manuscript, Christian-Albrechts-Universität, Kiel.

Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41-80). New York: Academic Press.

Frey, D. (1991). Informationssuche und Informationsbewertung bei Einzel- und Gruppenentscheidungen und mögliche Auswirkungen auf Politik und Wirtschaft. In D. Frey (Ed.), *37. Kongreß der Deutschen Gesellschaft für Psychologie, Kongreßband II* (pp. 45-56). Göttingen: Hogrefe.

Frey, D. & Stahlberg, D. (1986). Selection of information after receiving more or less reliable self-threatening information. *Personality and Social Psychology Bulletin, 12*, 434-441.

Frey, D., & Wicklund, R. (1978). A clarification of selective exposure: The impact of choice. *Journal of Experimental Social Psychology, 14*, 132-139.

Ganzach, Y. (1993). Goals as determinants of nonlinear noncompensatory judgment strategies: Leniency vs strictness. *Organizational Behavior and Human Decision Processes, 56*, 422-440.

Gibbins, M. (1984). Propositions about the psychology of professional judgment in public accounting. *Journal of Accounting Research, 22*, 103-125.

Gibbins, M. & Emby, C. (1985). Evidence on the nature of professional judgment in public accounting. *Auditing Research Symposium 1984*, 181-212.

Gilliland, S. W., & Schmitt, N. (1993). Information redundancy and decision behavior: A process tracing investigation. *Organizational Behavior and Human Decision Processes, 54*, 157-180.

Gilovich, T. (1983). Biased evaluation and persistence in gambling. *Journal of Personality and Social Psychology, 44*, 1110-1126.

Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology, 52*, 464-474.

Götz-Marchand, B., Götz, J., & Irle, M. (1974). Preference for dissonance reduction modes as a function of their order, familiarity, and reversibility. *European Journal of Social Psychology, 4*, 201-228.

Goldmann, K. (1971). International norms and governmental behavior. In K.Goldmann (Ed.), *International norms and war between states: Three studies in international politics* (Swedish studies in international relations) (pp.9-53). Stockholm: Läromodels förlagen.

Gordon, R. A., & Stuecher, U. (1993). The effect of anonymity and increased accountability on the linguistic complexity of teaching evaluations. *The Journal of Psychology, 126*, 639-649.

Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1988). The effect of applicant age, job level, and accountability on the evaluation of job applicants. *Organizational Behavior and Human Decision Processes, 41*, 20-33.

Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1989). The effect of applicant age, job level, and accountability on perceptions of female job applicants. *The Journal of Psychology, 123*, 59-68.

Grabitz, H.-J., & Grabitz-Gniech, G. (1973). Der kognitive Prozeß vor Entscheidungen: Theoretische Ansätze und experimentelle Untersuchungen. *Psychologische Beiträge, 15*, 522-549.

Grabitz, H.-J. & Haisch, J. (1972). Umbewertung von unterstützender und widersprechender Information vor Entscheidungen. *Archiv für Psychologie, 124*, 133-144.

Graumann, C. F. (1994). Verantwortung als soziales Konstrukt. *Zeitschrift für Sozialpsychologie, 25*, 184-191.

Greenfield, H. I. (1975). *Accountability in health facilities.* New York: Praeger.

Grush, J. E. (1978). Audiences can inhibit or facilitate competitive behavior. *Personality and Social Psychology Bulletin, 4*, 119-122.

Haccoun, R. R., & Klimoski, R. J. (1975). Negotiator status and accountability source: A study of negotiator behavior. *Organizational Behavior and Human Performance, 14*, 342-359.

Hagafors, R., & Brehmer, B. (1983). Does having to justify one's judgments change the nature of the judgment process? *Organizational Behavior and Human Performance, 31*, 223-232.

Hall, J. (1971). Decisions, decisions, decisions. *Psychology Today*, November 1971.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12*, 392-407.

Hamilton, V. L. (1978). Who is responsible? Toward a social psychology of responsibility attribution. *Social Psychology, 41*, 316-328.

Harkins, S. (1987). Social loafing and social facilitation. *Journal of Experimental Social Psychology, 23*, 1-18.

Harkins, S.G., & Jackson, J. M. (1985). The role of evaluation in eliminating social loafing. *Personality and Social Psychology Bulletin, 11*, 457-465.

Harkins, S. G., & Szymanski, K. (1987). Social loafing and social facilitation: New wine in old bottles. In C. Hendrick (Ed.), *Review of personality and social psychology 9: Group processes and intergroup relations* (pp. 167-188). Newbury Park: Sage.

Harkness, A. R., DeBono, K. G., & Borgida, E. (1985). Personal involvement and strategies for making contingency judgments: A stake in the dating game makes a difference. *Journal of Personality and Social Psychology, 49*, 22-32.

Hattrup, K., & Ford, J. K. (1995). The roles of information characteristics and accountability in moderating stereotype-driven processes during social decision making. *Organizational Behavior and Human Decision Processes, 63*, 73-86.

Hatvany, N., & Strack, F. (1980). The impact of a discredited key witness. *Journal of Applied Social Psychology, 10*, 490-509.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Heller, R.F., Saltzstein, H.D., Caspe, W.B. (1992). Heuristics in medical and non-medical decision making. *The Quarterly Journal of Experimental Psychology, 44A* (2), 211-235.

Hodgins, H. S., & Zuckerman, M. (1991). Beyond selecting information: Biases in spontaneous questions and resultant conclusions. *Journal of Experimental Social Psychology, 29*, 387-407.

Hogarth, R. M. (1987). *Judgment and choice* (2nd ed.). New York: Wiley.

Hollander, E. P. (1960). Competence and conformity in the acceptance of influence. *Journal of Abnormal and Social Psychology, 61*, 361-365.

Holton, B., & Pyszczynski, T. (1989). Biased information search in the interpersonal domain. *Personality and Social Psychology Bulletin, 15*, 42-51.

Hsee, C. K. (1995). Elastic justification: How tempting but task-irrelevant factors influence decisions. *Organizational Behavior and Human Decision Processes, 62*, 330-337.

Hsee, C. K. (1996). Elastic justification: How unjustifiable factors influence judgments. *Organizational Behavior and Human Decision Processes, 66*, 122-129.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9*, 90-98.

Huber, O. (1980).The influence of some task variables on cognitive operations in an information processing decision model. *Acta Psychologica, 45*, 187-196.

Huber, O. (1989). Information processing operators in decision making. In H. Montgomery & O. Svenson (Eds.), *Process and structure in decision making*. Chichester: Wiley.

Huber, O., & Seiser, G. (2001). Accounting and convincing: The effects of two types of justification on the decision process. *Journal of Behavioral Decision Making, 14*, 69-85.

Janis, I. L. (1958). *Psychological stress*. New York: Wiley.

Janis, I. L. (1959). Motivational factors in the resolution of decision conflicts. In M. R. Jones (Ed.), *Nebraska Symposium on Motivation* (Vol. 7). Lincoln: University of Nebraska Press.

Janis, I. L. (1982). *Groupthink: Psychological studies of policy decisions and fiascoes*. Boston, MA: Houghton Mifflin.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: Free Press.

Jecker, L. K. (1964).The cognitive effects of conflict and dissonance. In L. Festinger (Ed.), *Conflict, decision, and dissonance* (pp. 21-30). Stanford, CA: Stanford University Press.

Johnson, E. J., Meyer, R. J., & Ghose, S. (1989). When choice models fail: Compensatory representations in negatively correlated environments. *Journal of Marketing Research, 26*, 255-270

Johnson, V. E., & Kaplan, S.E. (1991). Experimental evidence on the effects of accountability on auditor judgments. *Auditing: A Journal of Practice & Theory, 10*, 96-107.

Jos, P. H., & Tompkins, M. E. (1994). Accountability: An interdisciplinary perspective. In J. L. Perry (Ed.), *Research in public administration* (pp. 41-104). Greenwich, CT: JAI Press.

Kahn, B. E., & Baron, J. (1995). An exploratory study of choice rules favoured for high-stakes decisions. *Journal of Consumer Psychology, 4*, 305-328.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80*, 237-251.

Kassin, S. M., Rigby, S., & Castillo, S. R. (1991). The accuracy-confidence correlation in eyewitness testimony: Limits and extensions of the retrospective self-awareness effect. *Journal of Personality and Social Psychology, 61*, 698-707.

Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York: Wiley.

Keeney, R., & Raiffa, H. (1993). *Decisions with multiple objectives: Preferences and value tradeoffs*. Cambridge: Cambridge University Press.

Kennedy, J. (1993). Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research, 31*, 231-245.

Kerr, N., & Bruun, S. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*, 78-94.

Kirby, S. L., & Davis, M. A. (1998). A study of escalating commitment in principal-agent relationships: Effects of monitoring and personal responsibility. *Journal of Applied Psychology, 83*, 206-217.

Klayman, J. (1983). Analysis of predecisional information search patterns. In P. C. Humphreys, O. Svenson, & A. Vari (Eds.), *Analyzing and aiding decision processes* (pp. 401-414). Amsterdam: North-Holland.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211-228.

Kleinhesselink, R. R. & Edwards, R. E. (1975). Seeking and avoiding belief-discrepant information as a function of its perceived refutability. *Journal of Personality and Social Psychology, 31*, 787-790.

Klimoski, R. J., & Ash, R. A. (1974). Accountability and negotiator behavior. *Organizational Behavior and Human Performance, 11*, 409-425.

Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Processes, 45*, 194-208.

Koele, P., & Westenberg, M. R. M. (1995). A compensation index for multiattribute decision strategies. *Psychonomic Bulletin and Review*, 2, 398-402.

Koonce, L., Anderson, U., & Marchant, G. (1995). Justification of decisions in auditing. *Journal of Accounting Research, 33, 369-384*.

Kramer, R. M., Pommerenke, P., & Newton, E. (1993). The social context of negotiation. Effects of social identity and interpersonal accountability on negotiator decision making. *Journal of Conflict Resolution, 37*, 633-654.

Kroon, M. B. R., van Kreveld, D., & Rabbie, J. M. (1991). Police intervention in riots: The role of accountability and group norms. A field experiment. *Journal of Community & Applied Social Psychology, 1*, 249-267.

Kroon, M. B. R., van Kreveld, D., & Rabbie, J. M. (1992). Group versus individual decision making. Effect of accountability and gender on groupthink. *Small Group Research, 23*, 427-458.

Kroon, M. B. R., t'Hart, P., & van Kreveld, D. (1991). Managing group decision making processes: Individual versus collective accountability and groupthink. *The International Journal of Conflict Management, 2*, 91-115.

Kruglanski, A. W. (1980). Lay epistemologic process and contents. *Psychological Review, 87*, 70-87.

Kruglanski, A. W. (1989). *Lay-epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum.

Kruglanski, A. W. (1990). Lay epistemic theory in social cognitive psychology. *Psychological Inquiry, 1*, 181-197.

Kruglanski, A. W. (1996). Motivated social cognition: Principles of the interface. In E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 493-520). New York: The Guilford Press.

Kruglanski, A. W., & Ajzen, I. (1983). Biases and error in human judgment. *European Journal of Social Psychology, 13*, 1-44.

Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impressional primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*, 448-468.

Kruglanski, A. W., & Klar, Y. (1987). A view from a bridge: Synthesizing the consistency and attribution paradigms from the lay epistemic perspective. *European Journal of Social Psychology, 17*, 211-241.

Kruglanski, A. W., & Meinholdt, C. (1990). Cognitive and motivational bases of judgmental bias: Toward a synthesis. *Polish Psychological Bulletin, 21*, 291-305.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53*, 636-647.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480-498.

Kunda, Z. (1999). *Social cognition: Making sense of people*. Cambridge, MA: MIT Press.

Larrick, R. P. (1993). Motivational factors in decision theories: The role of self-protection. *Psychological Bulletin, 113*, 440-450.

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.

Latané, B., Williams, K., & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology, 37*, 823-832.

Lazarus, R. S. (1983). The costs and benefits of denial. In S. Breznitz (Ed.), *The denial of stress* (pp. 1-30). Madison, CT: International Universities Press.

Lee, H., Herr, P. M., Kardes, F. R., & Kim, C. (1999). Motivated search: Effects of choice accountability, issue involvement, and prior knowledge on information acquisition and use. *Journal of Business Research, 45*, 75-88.

Lerner, J. S. (1994). *Accountability for allocation: Do one thing but say another. Unpublished master's thesis*. University of California, Berkeley.

Lerner, J. S., Goldsmith, J. H., & Tetlock, P. E. (1998). Sober second thought: The effect of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin, 24*, 563-574.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*, 255-275

Lewin, K. (1938). The conceptual representation and measurement of psychological forces. *Contributions to Psychological Theory, 1*.

Liberman, A. & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin, 18*, 669-679.

Lichtenstein, S., Gregory, R., Slovic, P., & Wagenaar, W. A. (1990). When lives are in your hands: Dilemmas of the societal decision maker. In R.M. Hogarth (Ed.), *Insights in decision making* (pp. 91-106). Chicago: University of Chicago Press.

Lipe, M. G. (1982). *A cross-study analysis of covariation judgments*. Working Paper No. 96, University of Chicago.

Lord, A. T. (1992). Pressure: A methodological consideration for behavior research in auditing. *Auditing: A Journal of Practice & Theory, 11*, 89-108.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.

Lowin, A. (1967). Approach and avoidance: Alternative modes of selective exposure to information. *Journal of Personality and Social Psychology, 6*, 1-9.

Lowin, A. (1969). Further evidence for an approach-avoidance interpretation of selective exposure. *Journal of Experimental Social Psychology, 5*, 265-271.

McCloskey; M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114*, 3-18.

Magjuka, R. J., Baldwin, T. T., & Loher, B. T. (1994). The combined effects of three pretraining strategies on motivation and performance: An empirical exploration. *Journal of Management Issues, 3*, 282-296.

Malmsten, N. (1993). *Attractiveness and/or importance restructuring in a choice with environmental related attributes*. Report No. 40, Cognition and Decision Research Unit, University of Stockholm.

Mandel, D. R., Axelrod, L. J., & Lehman, D. R. (1993). Integrative complexity in reasoning about the Persian Gulf War and the accountability to a skeptical audience hypothesis. *Journal of Social Issues, 49*, 201-215.

Markus, H. (1978). The effect of mere presence on social facilitation: An unobtrusive test. *Journal of Experimental Social Psychology, 14*, 389-397.

McAllister, D. W., Mitchell, T. R., & Beach, L. R. (1979). The contingency model for the selection of decision strategies: An empirical test of the effects of significance, accountability, and reversibility. *Organizational Behavior and Human Performance, 24*, 228-244.

McKenna, F. P., & Myers, L. B. (1997). Illusory self-assessments - can they be reduced? *British Journal of Psychology, 88*, 39-51.

McLean Parks, J., & Conlon E. J. (1990). Justification and the processing of information. *Journal of Applied Social Psychology, 20*, 703-723.

McKeon, R. (1957). The development and the significance of the concept of responsibility. *Revue Internationale de Philosophie, 39*, 3-32.

Mero, N. P., & Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology, 80*, 517-524.

Messier, W. F., Jr., & Quilliam W.C. (1992). The effect of accountability on judgment: Development of hypotheses for auditing. *Auditing: A Journal of Practice & Theory, 11*, 123-138.

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology, 83*, 340-363.

Miller, D. T., & Ross, M. (1975). Self-serving biases in attribution of causality: Fact or fiction? *Psychological Bulletin*, *82*, 213-225.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart, & Winston.

Mills, C. W. (1940). Situated actions and vocabularies of motive. *American Sociological Review, 5*, 904-913.

Mintzberg, H. (1989). *Mintzberg on management: Inside our strange world of organizations*. New York: Free Press.

Montgomery, H. (1983). Decision rules and the search for a dominance structure: Towards a process model of decision making. In P. Humphreys, O. Svenson, & A. Vari (Eds.), *Analyzing and aiding decision processes* (pp. 343-369). Amsterdam: North Holland.

Montgomery, H. (1989). From cognition to action: The search for dominance in decision making. In H. Montgomery & O. Svenson (Eds.), *Process and structure in decision making* (pp. 23-49).New York: Wiley.

Mosier, K. L., Skitka, L., Heers, S., & Burdick, M. (1998). Automation bias: Decision making and performance in high-tech cockpits. *The International Journal of Aviation Psychology, 8*, 47-63.

Murphy, R. (1994). The effect of task characteristics on covariation assessment: The impact of accountability and judgment frame. *Organizational Behavior and Human Decision Processes, 60*, 139-155.

Newman, E., & Turem, J. (1974). The crisis of accountability. *Social Work*, *January 1974*, 5-16.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Nisbett, R. E., Zukier, H., & Lemley, R. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology, 13*, 248-277.

Olson, J. M., & Zanna, M. P. (1979). A new look at selective exposure. *Journal of Experimental Social Psychology, 15*, 1-15.

Ordonez, L. D., Benson III, L., & Beach, L. R. (1999). Testing the compatibility test: How instructions, accountability, and anticipated regret affect prechoice screening of options. *Organizational Behavior and Human Decision Processes, 78*, 63-80.

O'Neill, P. (1989). Responsible to whom? Responsible for what? Some ethical issues in community intervention. *American Journal of Community Psychology*, *17*, 323-341.

O'Reilly, C. A. (1983). The use of information in organizational decision making: A model and some propositions. *Research in Organizational Behavior, 5*, 103-139.

Oskamp, S. (1965). Overconfidence in case study judgments. *Journal of Consulting Psychology, 29*, 261-165.

Oswald, M. E. (1993). Hypothesentesten: Suche und Verarbeitung hypothesenkonformer und hypothesenkonträrer Informationen. In W. Hell, K. Fiedler, & G. Gigerenzer (Eds.), *Kognitive Täuschungen: Fehlleistungen und Mechanismen des Urteilens, Denkens und Erinnerns* (pp. 189-211). Heidelberg: Spektrum Akademischer Verlag.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*, 366-387.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin, 92*, 382-402.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. Cambridge: University Press.

Peecher, M. E., & Kleinmuntz, D. N. (1991). Accountability and justification in audit judgment processes: A comment on Johnson and Kaplan. *Auditing: A Journal of Practice & Theory*, *10*, 108-113.

Pelham, B. W., & Neter, E. (1995). The effect of motivation of judgment depends on the difficulty of judgment. *Journal of Personality and Social Psychology, 68*, 581-594.

Pennington, J., & Schlenker, B. R. (1999). Accountability for consequential decisions: Justifying ethical judgments to audiences. *Personality and Social Psychology Bulletin, 25*, 1067-1081.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches*. Dubuque, IL: William C. Brown.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 19, pp. 123-205). San Diego, CA: Academic Press.

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research, 10*, 135-146.

Pfeffer, J. (1981). Management as symbolic action: The creation and maintenance of organizational paradigms. *Organizational Behavior, 3*, 1-52.

Pohl, R. F. (1995). Disenchanting hindsight bias. In J.-P. Caverni, M. Bar-Hillel, F.H. Barron, & H. Jungermann (Eds.), *Contributions to Decision-Making I* (pp. 323-334). Amsterdam: North-Holland.

Pratt, J. W., & Zeckhauser, R. J. (1985). Principals and agents: An overview. In J. W. Pratt & R. J. Zeckhauser, *Principals and agents: The structure of business*. Boston: Harvard Business School Press.

Prentice-Dunn, S., & Rogers, R. W. (1982). Effects of public and private self-awareness on deindividuation and aggression. *Journal of Personality and Social Psychology, 43*, 503-513.

Pruitt, D. G. (1981). *Negotiation behavior*. New York: Academic Press.

Pruitt; D.G., & Carnevale, P. J. D. (1982). The development of integrative agreements. In V. J. Derlega & J. Grzelak (Eds.), *Cooperation and helping behavior: Theories and research*. New York: Academic Press.

Pruitt, D. G., Kimmel, M. J., Britton, S., Carnevale, P. J. D., Magenau, J. M., Peragallo, J., & Engram, P.(1978). The effect of accountability and surveillance on integrative bargaining. In H. Sauermann (Ed.), *Bargaining behavior* (pp. 310-343). Tübingen: J. C. B. Mohr (Paul Siebeck).

Pyszczynski, T. & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis testing model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 297-340). New York: Academic Press.

Pyszczynski, T., Greenberg, J., & Holt, K. (1985). Maintaining consistency between self-serving beliefs and available data: A bias in information evaluation. *Personality and Social Psychology Bulletin, 11*, 179-190.

Raiffa, H. (1969). *Decision analysis*. Reading, MA: Addison-Wesley.

Ranyard, R. (1991). Structure and strategy in justifying environmental decisions. *Journal of Environmental Psychology, 11*, 43-57.

Reicher, S. D. (1987). Crowd behaviour as social action. In J.C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 171-202). Oxford: Blackwell.

Roberts, K. H., Stout, S. K., & Halpern, J. J. (1994). Decision dynamics in two high reliability military organizations. *Management Science, 40*, 614-624.

Roloff, M. E., & Campion, D. E. (1987). On alleviating the debilitating effects of accountability on bargaining: Authority and self-monitoring. *Communication Monographs, 54*, 145-164.

Romzek, B. S., & Dubnick, M. J. (1987). Accountability in the public sector: Lessons from the challenger tragedy. *Public Administration Review, 37*, 227-238.

Rosenberg, M.L., & Brody, R. (1974). The threat or challenge of accountability. Social work, 19, 344-350.

Ross, A. (1975). *On guilt, responsibility and punishment*. London: Steven & Sons.

Ross, C., & Joelson, J. B. (1986). Holding accountability systems accountable: Is what's good for the goose good for the gander? *The Journal of Applied Social Sciences*, *10*, 156-175.

Ross, L. (1977). The intuitive psychologist and his shortcomings. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). San Diego, CA: Academic Press.

Ross, L., Greene, D., & House, P. (1977). The false consensus phenomenon: An attributional bias in self-perception and social perception processes. *Journal Of Experimental Social Psychology, 13*, 279-301.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self perception and social perception processes. *Journal of Personality and Social Psychology, 32*, 880-892.

Rozelle, R. M., & Baxter, J. C. (1981). Influence of role pressures on the perceiver: Judgments of videotaped interviews varying judge accountability and responsibility. *Journal of Applied Psychology, 66*, 437-441.

Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 676-696.

Russo, J. E., Medvec, V. H., & Meloy, M. G. (1996). The distortion of information during decisions. *Organizational Behavior and Human Decision Processes, 66*, 102-110.

Sanbonmatsu, D. M., Akimoto, S. A., & Biggs, E. (1993). Overestimating causality: Attributional effects of confirmatory processing. *Journal of Personality and Social Psychology, 65*, 892-903.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*, 229-241.

Schadewald, M. S., & Limberg, S. T. (1992). Effect of information order and accountability on causal judgments in legal context. *Psychological Reports, 71*, 619-625.

Schaller, M. (1992). In-group favoritism and statistical reasoning in social inference: Implications for formation and maintenance of group stereotypes. *Journal of Personality and Social Psychology, 63*, 61-74

Schlenker, B. R. (1980). *Impression management: The self concept, social identity, and interpersonal relations*. Monterey, CA: Brooks/Cole.

Schlenker, B. R. (1982). Translating actions into attitudes: An identity-analytic approach to the explanation of social conduct. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 15, pp. 193-247). New York: Academic Press.

Schlenker, B. R. (1986). Self-identification: Toward an integration of the private and public self. In R.F. Baumeister (Ed.), *Public self and private self* (pp. 21-55). New York: Springer.

Schlenker, B. R., Britt, T. W., Pennington, J., Murphy, R., & Doherty, K. (1994). The triangle model of responsibility. *Psychological Review, 101*, 632-652.

Schlenker, B. R., & Weigold, M. F. (1989). Self-identification and accountability. In A. G. Giacalone & P. Rosenfeld (Eds.), *Impression management in the organization* (pp. 21-43). Hillsdale: Lawrence Erlbaum.

Schlenker, B. R., Weigold, M. F., & Doherty, K. (1991). Coping with accountability: Self-identification and evaluative reckonings. In C. R. Snyder & D. R. Forsyth (Eds.), *Handbook of social and clinical psychology: The health perspective* (pp. 96-115). New York: Pergamon.

Schoenrade, P. A., Batson, C. D., Brandt, J. R., & Loud, Jr., R. E. (1986). Attachment, accountability, and motivation to benefit another in distress. *Journal of Personality and Social Psychology, 51*, 557-563.

Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference: *Journal of Experimental Psychology: General, 110*, 101-120.

Schwarz, N., Frey, D., & Kumpf, M. (1980). Interactive effects of writing and reading a persuasive essay on attitude change and selective exposure. *Journal of Experimental Social Psychology, 16*, 1-17.

Schwartz, S. H. (1977). Normative influences on altruism. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 221-279). New York: Academic Press.

Sears, D. O. (1965). Biased indoctrination and selectivity of exposure to new information. *Sociometry, 28*, 420-426.

Selart, M. (1996). Structure compatibility and restructuring in judgment and choice. *Organizational Behavior and Human Decision Processes, 65*, 106-116.

Semin, G. R., & Manstead, A. S. R. (1983*). The accountability of conduct: A social psychological analysis*. London: Academic Press.

Shafir, E. (1993). Choosing vs. rejecting: Why some options are both better and worse than others. *Memory & Cognition, 21*, 546-556.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*,11-36.

Shaver, K. G. (1985). *The attribution of blame: Causality, responsibility, and blameworthiness*. New York: Springer Verlag

Sheer, V. C., & Weigold, M. F. (1995). Managing threats to identity: The accountability triangle and strategic accounting. *Communication Research, 22*, 592-611.

Sherman, B. R., & Kunda, Z. (1989, June). *Motivated evaluation of scientific evidence*. Paper presented at the American Psychological Society convention, Arlington.

Showers, C., & Cantor, N. (1985). Social cognition: A look at motivated strategies. *Annual Review of Psychology, 36*, 275-305.

Siegel-Jacobs, K., & Yates, J. F. (1996). Effects of procedural and outcome accountability on judgment quality. *Organizational Behavior and Human Decision Processes, 65*, 1-17.

Simon, H. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99-118.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16*, 158-174.

Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of Consumer Research, 19*, 105-118.

Simonson, I., & Nye, P. (1992). The effect of accountability on susceptibility to decision errors. *Organizational Behavior and Human Decision Processes, 51*, 416-446.

Simonson, I., & Staw, B. M. (1992). Deescalation strategies: A comparison of techniques for reducing commitment to losing courses of action. *Journal of Applied Psychology, 77*, 419-426.

Skov, R.B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22*, 93-121.

Slovic, P. (1975). Choice between equally valued alternatives. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 280-287.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1988). Response mode, framing, and information-processing effects in risk assessment. In D.E. Bell, H. Raiffa & A. Tversky (Eds.), *Decision making: Descriptive, normative and prescriptive interactions* (pp. 152-166). Cambridge: Cambridge University Press.

Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological Bulletin, 109*, 472-489.

Snyder, M., & Campbell, B. (1980). Testing hypotheses about other people: The role of the hypothesis. *Personality and Social Psychology Bulletin, 6*, 421-426.

Snyder, M., & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. *Journal of Personality and Social Psychology, 15*, 330-342.

Snyder, M., & Swann, W. B., Jr. (1978a). Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology, 14*, 148-162.

Snyder, M., & Swann, W. B., Jr. (1978b). Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology, 36*, 1201-1212.

Snyder, M. & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology, 36*, 941-950.

Snyder, M. L., Kleck, R. E., Strenta, A., & Mentzer, S. J. (1979). Avoidance of the handicapped: An attributional ambiguity analysis. *Journal of Personality and Social Psychology, 37*, 2297-2306.

Staub, E. (1978). *Positive social behavior and morality* (Vol.1, Social and personal influences). New York: Academic Press.

Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16*, 27-44.

Staw, B. M. (1980). Rationality and justification in organizational life. *Research in Organizational Behavior, 2*, 45-80.

Staw, B. M., & Boettger, R. D. (1990). Task revision: A neglected form of work performance. *Academy of Management Journal, 33*, 534-559.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 261-302). New York: Academic Press.

Steele, C. M., & Liu, T. J. (1981). Making the dissonance act unreflective of the self: Dissonance avoidance and the expectancy of a value affirming response. *Personality and Social Psychology Bulletin, 7*, 383-387.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology, 45*, 5-19.

Stewart, D. D., Billings, R. S., & Stasser, G. (1998). Accountability and the discussion of unshared, critical information in decision-making groups. *Group Dynamics: Theory, Research, and Practice, 2*, 18-23.

Stokmans, M. J. W. (1991). *The relative importance of product attributes: Consumer theories in new-product development*. Delft: Delft University Press.

Stone, D. N. & Schkade, D. A. (1991). Numeric and linguistic information representation in multiattribute choice. *Organizational Behavior and Human Decision Processes, 49*, 42-59.

Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. Martin & A. Tesser (Eds.), *The construction of social judgment* (pp. 249-275). Hillsdale, NJ: Lawrence Erlbaum.

Strohmer, D. C., & Shivy, V. A. (1994). Bias in counselor hypothesis testing: Testing the robustness of counselor confirmatory bias. *Journal of Counseling & Development, 73*, 191-197.

Svenson, O. (1979). Process descriptions of decision making. *Organizational Behaviour and Human Performance, 23*, 86-112.

Svenson, O. (1990). Some propositions for the classification of decision situations. In K. Borcherding, O. I. Larichev, & D. M. Messick (Eds.), *Contemporary issues in decision making*. North-Holland: Elsevier.

Svenson, O. (1992). Differentiation and consolidation theory of human decision making: A frame of reference for the study of pre- and post-decision processes. *Acta Psychologica, 80*, 143-168.

Svenson, O. (1996). Decision making and the search for fundamental psychological regularities: What can be learned from a process perspective. *Organizational Behavior and Human Decision Processes, 65*, 252-267.

Svenson, O., & Benthorn, L. J. (1992). Consolidation processes in decision making: Post-decision changes in attractiveness of alternatives. *Journal of Economic Psychology, 13*, 315-327.

Svenson, O., & Hill, T. (1996). Turning prior disadvantages into advantages: Differentiation and consolidation in real-life decision making. In R. Ranyard, W. R. Crozier, & O. Svenson (Eds.), *Decision making: Cognitive models and explanations*. London: Routledge.

Svenson, O., & Malmsten, N. (1991). Post decision consolidation processes in human decision making. *Comunicazioni Scientifiche di Psycologia Generale, 6*, 109-120.

Svenson, O., & Malmsten, N. (1996). Post-decision consolidation over time as a function of gain or loss of an alternative. *Scandinavian Journal of Psychology, 37*, 302-311.

Svenson, O., & Shamoun, S. (1997). Predecision conflict and different patterns of postdecision attractiveness restructuring: Empirical illustrations from an important real-life decision. *Scandinavian Journal of Psychology, 38*, 243-251.

Svenson, O., Ortega Rayo, A., Andersen, M., Sandberg, A., & Svahlin, I. (1994). Post-decision consolidation, as a function of the instructions to the decision maker and of the decision problem. *Acta Psychologica, 87*, 181-197.

Sweeney, P. D., & Gruber, K. L. (1984). Selective exposure: Voter information preferences and the Watergate affair. *Journal of Personality and Social Psychology, 46*, 1208-1221.

Takemura, K. (1993). The effect of decision frame and decision justification on risky choice. *Japanese Psychological Research, 35*, 36-40.

Tausky, C., & Chelte, A. F. (1983). Accountability and productivity: Some longitudinal data. *Work and Occupations, 10*, 207-220.

Taylor, K. A. (1995). Testing credit and blame attributions as explanations for choices under ambiguity. *Organizational Behavior and Human Decision Processes, 64*, 128-137.

Taylor, S. E. (1981). The interface of cognitive and social psychology. In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 189-212). Hillsdale, NJ: Lawrence Erlbaum.

Tedeschi, J. T. (Ed.) (1981). *Impression management theory and social psychological research*. New York: Academic Press.

Tedeschi, J. T., & Riess, M. (1981). Predicaments and verbal tactics of impression management. In C. Antaki (Ed.), *Ordinary language explanations of social behavior* (pp. 271-309). London: Academic Press.

Tetlock, P. E. (1983a). Accountability and complexity of thought. *Journal of Personality and Social Psychology, 45*, 74-83.

Tetlock, P. E. (1983b). Accountability and the perseverance of first impressions. *Social Psychology Quarterly, 46*, 285 - 292.

Tetlock, P. E. (1985a). Accountability: The neglected social context of judgment and choice. In B. M. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 7, pp. 297-332). Greenwich, CT: JAI Press.

Tetlock, P. E. (1985b). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly, 48*, 227- 236.

Tetlock, P. E. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology, 52*, 700-709.

Tetlock, P.E. (1991). An alternative metaphor in the study of judgment and choice: People as politicians. *Theory & Psychology, 1*, 451-475.

Tetlock, P.E. (1992). The impact of accountability on judgment and choice: Toward a social contingency model. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 331-376). San Diego: Academic Press.

Tetlock, P. E., & Boettger, R. (1989). Accountability: A social magnifier of the dilution effect. *Journal of Personality and Social Psychology, 57*, 388-398.

Tetlock, P. E., & Boettger, R. (1994). Accountability amplifies the status quo effect when change creates victims. *Journal of Behavioral Decision Making, 7*, 1-23.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgement processes in a personality prediction task. *Journal of Personality and Social Psychology, 52*, 700-709.

Tetlock, P. E., & Levi, A. (1982). Attribution bias: On the inconclusiveness of the cognition-motivation debate. *Journal of Experimental Social Psychology, 18*, 68-88.

Tetlock, P. E., Lerner, J. S., & Boettger, R. (1996). The dilution effect: Judgmental bias, conversational convention, or a bit of both? *European Journal of Social Psychology, 26*, 915-934.

Tetlock, P.E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: conformity, complexity, and bolstering. *Journal of Personality and Social Psychology, 57*, 632-640.

Thompson, L. (1995). They saw a negotiation: Partisanship and involvement. *Journal of Personality and Social Psychology, 68*, 839-853.

Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology, 43*, 22-34.

Trope, Y., & Bassok, M. (1983). Information-gathering strategies in hypothesis-testing. *Journal of Experimental Social Psychology, 19*, 560-576.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*, 281-299.

Tyszka, T. (1983). Contextual multi-attribute decision rules. In L. Sjöberg, T. Tyszka, & J. Wise (Eds.), *Human decision making*. Lund: Doxa.

Van Hiel, A., & Schittekatte, M. (1997). Information exchange in context: Effects of gender composition of group, accountability, and intergroup perception on group decision making. *Journal of Applied Social Psychology, 28*, 2049-2067.

Von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge: Cambridge University Press

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.

Webster, D. M., Richter, L., & Kruglanski, A. W. (1996). On leaping to conclusions when feeling tired: Mental fatigue effects on impressional primacy. *Journal of Experimental Social Psychology, 32*, 181-195.

Weigold, M. F., & Schlenker, B. R. (1991). Accountability and risk taking. *Personality and Social Psychology Bulletin, 17*, 25-29.

Weiner, B. (1995a). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: The Guildford Press.

Weiner, B. (1995b). Inferences of responsibility and social motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 1-47). San Diego: Academic Press.

Weldon, E., & Gargano, G. M. (1985). Cognitive effort in additive task groups: The effects of shared responsibility on the quality of multiattribute judgments. *Organizational Behavior and Human Decision Processes, 36*, 348-361.

Weldon, E., & Gargano, G. M. (1988). Cognitive loafing: The effects of accountability and shared responsibility on cognitive effort. *Personality and Social Psychology Bulletin, 14*, 159-171.

Wells, G. L., Petty, R. E., Harkins, S. G., Kagehiro, D., & Harvey, J. (1977). Anticipated discussion of interpretation eliminates actor-observer differences in the attribution of causality. *Sociometry, 40*, 247-253,

Wernicke, J. (1993). *Consolidation processes and post-decision dissonance reduction*. Unpublished diploma thesis, Technische Hochschule Darmstadt.

Westenberg, M. R. M. (1991). *Response modes and decision strategies*. Doctoral dissertation, University of Amsterdam.

Westenberg, M. R. M., & Koele, P. (1992). Response modes, decision processes, and decision outcomes. *Acta Psychologica, 80*, 169-184.

Wicklund, R.A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Hillsdale, NJ: Lawrence Erlbaum.

Williams, K., Harkins, S. G., & Latané, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40*, 303-311.

Wyer, R. S., & Frey, D. (1983). The effects of feedback about self and others on the recall and judgments of feedback-relevant information. *Journal of Experimental Social Psychology, 19*, 540-559.

Yarnold, P. R., Mueser, K. T., & Lyons, J. S. (1988). Type A behavior, accountability, and work rate in small groups. *Journal of Research in Personality, 22*, 353-360.

Zajonc, R.B. (1965). Social facilitation. *Science, 149*, 269-274.

Zimbardo, P.G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W.J. Arnold & D. Levine (Eds.), *Nebraska Symposium on Motivation* (Vol. 17, pp. 237-307). Lincoln: University of Nebraska Press.

Zukier, H. (1982). The dilution effect: The role of correlation and the dispersion of predictor variables in the use of non-diagnostic information. *Journal of Personality and Social Psychology, 43*, 1163-1174.

APPENDICES

Appendix 1. Information set 1, experiment 1.

	A	**B**	**C**	**D**
Diploma thesis grade	B	A-	A	B+
Educational stay in a foreign country	none	3 months placement in the USA	studied in England for one year	6 weeks in England on a school exchange
Number of placements in the area of Org. Psych.	3, in different fields	none	1	2, in similar fields
Job as a student research assistant	for one year, when writing thesis	demonstrator in a lab class for one term	none	for several years, since intermediate exams
Activities as a student member of a professional organisation	helping to create a data base for placements	member of BDP, active as a student rep and in EFPSA	member of BDP	no activities

Appendix 2. Information set 2, experiment 1.

	A	**B**	**C**	**D**
Computing skills	good knowledge of 1 word processing package and SPSS	medium knowledge of 1 word processing package and SPSS	fair knowledge of 1 word processing package and SPSS	very good knowledge of 2 word processing packages, SPSS, and a programming language, experience with data bases
Knowledge of foreign languages	English: fair no French	English: very good French: good	English: excellent French: very good	English: good French: fair
Activities as a student rep on university committees	two years on Departmental Board	AstA member for several years	substitute on Faculty Board for one term	no
Relevance of thesis topic to Org. Psych.	thesis in Organisational Psychology, in research area of future supervisor	thesis in an area unrelated to Organisational Psychology	thesis in Clinical Psychology with possible applications in Org. Psych.	thesis in Organisational Psychology, not in research area of future supervisor
Reference from thesis supervisor	fair	very good	extremely good	good

Appendix 3. Mean values (*M*) and standard deviations (*SD*) for ratings of similarity between attribute pairs in experiment 2 (ratings for parallel attributes appear in bold typeface).

Similarity between	M	SD
user interface design & satisfaction with user interface design	**4.82**	0.55
user interface design & satisfaction with method support	2.72	1.15
user interface design & method support	2.54	1.17
user interface design & training requirements	3.30	1.36
user interface design & training costs	2.92	1.39
user interface design & hardware costs	2.37	1.12
user interface design & hardware requirements	3.00	1.18
satisfaction with user interface design & satisfaction with method support	2.84	1.20
satisfaction with user interface design & method support	2.74	1.28
satisfaction with user interface design & training requirements	2.84	1.30
satisfaction with user interface design & training costs	2.98	1.35
satisfaction with user interface design & hardware costs	2.27	1.21
satisfaction with user interface design & hardware requirements	2.56	1.17
satisfaction with method support & method support	**4.70**	0.79
satisfaction with method support & training requirements	2.80	1.02
satisfaction with method support & training costs	2.52	1.23
satisfaction with method support & hardware costs	1.85	1.02
satisfaction with method support & hardware requirements	2.02	0.94
method support & training requirements	2.86	1.17
method support & training costs	2.79	1.17
method support & hardware costs	2.06	1.09
method support & hardware requirements	2.37	1.18
training requirements & training costs	**4.35**	1.09
training requirements & hardware costs	1.37	0.70
training requirements & hardware requirements	1.54	0.98
training costs & hardware costs	1.64	1.11
training costs & hardware requirements	1.56	0.81
hardware costs & hardware requirements	**4.76**	0.60

Note: Similarity ratings were given on a scale from 1 (very dissimilar) to 5 (very similar).

Appendix 4. Information set 1, experiment 2.

	A	**B**	**C**	**D**
Satisfaction with Method Support	judgement of testers: 6	judgement of testers: 7	judgement of testers: 3	judgement of testers: 4
Training Requirements	new training concept needed, 25 days for development, 5-day course	new training concept needed, 20 days for development, 5-day course	adaptation of vendor's concept, 2 days for development, 5-day course	adaptation of vendor's concept, 4 days for development, 5-day course
Hardware Costs	ca. 200.000 DM	ca. 340.000 DM	ca. 230.000 DM	ca. 300.000 DM
User Interface Design	medium-suited to requirements, medium orientation, flexible user guidance, no context-sensitive help	well-suited to requirements, medium orientation, flexible user guidance, context-sensitive help	medium-suited to requirements, good orientation, inflexible user guidance, context-sensitive help	well-suited to requirements, good orientation, flexible user guidance, context-sensitive help

Appendix 5. Information set 2, experiment 2.

	A	B	C	D
Hardware Requirements	8 MB RAM, 20 MB on hard disk, 486/33 MHz processor, laser printer, 16'' monitor	16 MB RAM, 150 MB on hard disk, Pentium/90 MHz processor, postscript laser printer, 19'' monitor	12 MB RAM, 50 MB on hard disk, 486/66 MHz processor, laser printer, 16'' monitor	16 MB RAM, 120 MB on hard disk, 486/100 MHz processor, laser printer, 19'' monitor
Satisfaction with User Interface Design	judgement of testers: 2	judgement of testers: 5	judgement of testers: 3	judgement of testers: 6
Method Support	entity-relationship and data flow modelling, modelling of organisational procedures has to be configured manually	entity-relationship modelling, data flow modelling and modelling of organisational procedures	entity-relationship and data flow modelling, no modelling of organisational procedures	entity-relationship modelling and modelling of organisational procedures, no data flow modelling
Training Costs	ca. 50.000 DM	ca. 45.000 DM	ca. 7.000 DM	ca. 9.000 DM